DOCUMENT RESUME

ED 043 676 TM 000 141

TITLE	Report on the Second Florida Instructional Seminar on Evaluation.
INSTITUTION	Florida State Dept. of Education, Tallahassee.; Orange County Board of Public Instruction, Orlando, Fla.
PUB DATE	Aug 68
GRANT	OEG-68-06745-7
NOTE	77p.; Presented at the Second Florida Instructional Seminar on Evaluation, Orlando, Florida, August 1968
EDRS PRICE	EDRS Price MF-$0.50 HC-$3.95
DESCRIPTORS	Counseling Programs, Educational Programs, Evaluation Criteria, *Evaluation Techniques, Guidance Programs, *Instructional Programs, *Models, *Program Design, Program Effectiveness, *Program Evaluation, Public Schools, Reading Programs, School Personnel
IDENTIFIERS	*Florida Instructional Program on Evaluation

ABSTRACT

The participants in the Second Florida Seminar on Evaluation were divided into two principal groups: Technical Evaluation Information and Evaluation Models in Selected Areas. In the Technical Evaluation Information subgroups, technical material involved in the evaluation process was presented. A topical outline of this material is provided. The Evaluation Models subgroups developed models in the areas of reading (two models), communication skills, arithmetic, guidance, and self-concept programs. These models are part of the report. The texts of two of the major addresses are included: "Evaluation and the Two-Party Monologue" by Jarvis Barnes and "Education for Imagination" by Harry M. Sparks. Finally, the seminar itself is evaluated and suggestions and recommendations for follow-up programs are presented. (PR)

ED043676

REPORT

on the

SECOND FLORIDA INSTRUCTIONAL SEMINAR ON EVALUATION

13-16 AUGUST 1968

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED043676

REPORT

on the

SECOND FLORIDA INSTRUCTIONAL SEMINAR ON EVALUATION

Held in

ORLANDO, FLORIDA

at the

ROBERT MEYER MOTOR INN

on

13-16 AUGUST 1968

Sponsored by

ORANGE COUNTY BOARD OF PUBLIC INSTRUCTION

and

FLORIDA STATE DEPARTMENT OF EDUCATION

TM 000 141

Preface

The first Florida Instructional Seminar on Evaluation was held in Orlando on August 1-4, 1967, as a result of a felt need among State Department of Education and county level personnel for assistance in evaluating educational programs, with particular emphasis on those approved under the Elementary and Secondary Education Act of 1965. Evaluation of this first Seminar indicated a strong need for a second Seminar to assist the public school personnel to become more knowledgeable and to do a better job of evaluating school instructional programs. This need was intensified by the requirements of the new Florida law passed by the Legislature in the spring of 1968 which provided assistance for instructional programs under the provisions of the Educational Improvement Expense (EIE) fund. Programs financed through these funds have to be evaluated and reports made to the State Department of Education.

Another Mini-Grant was filed early in 1968 to secure funds for conducting a follow-up Seminar during the summer. Since it was not approved until July 1, the earliest possible dates for conducting the Seminar were August 13-16.

TABLE OF CONTENTS

CHAPTER		PAGE
	PREFACE	ii
I.	DESCRIPTION OF THE SEMINAR PROGRAM	1
II.	ADDRESSES OF SPEAKERS	4
	Evaluation and the Two-Party Monologue — Jarvis Barnes	4
	Education for Imagination — Harry M. Sparks	14
III.	AREA A — EVALUATIVE MODELS IN SELECTED AREAS	20
	Introduction	20
	Model for Reading Program I — Staff Development: Visual Perception	21
	Model for Reading Program II	24
	Model for Arithmetic — A Comparison of Two Methods of Teaching Regrouping in Addition at the Primary Level	26
	Model for Communication Skills Program — Suggested Behavioral Objectives for One Area in a Communication Skills Program	29
	Model for Guidance Program — Guidance	34
	The Effects of Group Counseling on Under-achieving Ninth Grade Students	37
	Model for Self-Concept Program — A Project to Improve Self-Concept of Disadvantaged Second Grade Children	39
IV.	AREA B — TECHNICAL EVALUATION INFORMATION	47
	Introduction	47
	Tentative Topical Outline — Area B	47
	Evaluation — Area B	51
V.	SEMINAR EVALUATION	63
	Evaluation of Seminar Objectives	63
	Evaluation of Individual Objectives	65
	General Comments and Suggestions	66
VI.	RECOMMENDATIONS FOR FOLLOW-UP	68
	APPENDIX — SEMINAR LEADERSHIP PERSONNEL	70

CHAPTER I

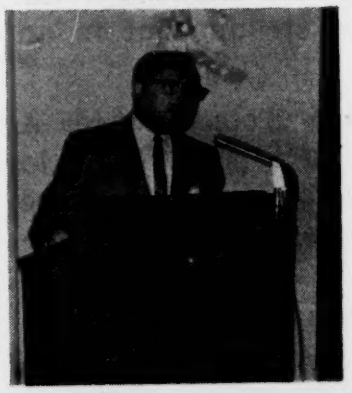
Description of The Seminar Program

The specific purpose of this second Seminar was to follow-up the work accomplished the preceding summer and to provide opportunities for school personnel to improve their skills in designing instructional programs in such a way that they could be properly evaluated. Special emphasis was given to the Federal programs fianced under Federal grants.

The State-wide Planning Committee decided again to invite two representatives from each of Florida's sixty-seven counties, two from each of the seven state-supported universities and two from each of the four Catholic dioceses.

The participants were divided into two principal areas: Area A — Evaluative Models in Selected Areas, and Area B — Technical Evaluation Information. Dr. J. B. White, one of the General Consultants, was charged with the planning and direction of activities in Area A, and Dr. Hazen A. Curtis, the other General Consultant, was charged with the same responsibility in Area B. This procedure enabled the planning and direction to take a more definite course than the previous year.

The participants in each of the two areas were divided further into six working groups. Outstanding consultants were then selected to work directly with each of these twelve working groups. During the period of the Seminar these working groups met for morning and afternoon sessions.



At noon on each of the four days, the participants met for a luncheon and a general session. The keynote address entitled "Evaluation And The Two-Party Monologue" was given by Dr. Jarvis Barnes, Assistant Superintendent for Research and Development for the Atlanta Public Schools. He stressed four components of evaluation:

(a) the status evaluation has attained
(b) a few principles of evaluation
(c) evaluation as a subsystem within a school system, and
(d) cost effectiveness.

At the second general session on Wednesday, Richard Y. Jaeger, from the Program Planning and Evaluation Branch, U.S. Office of Education, appraised the group of the present status of Federal planning and evaluation in Washington. Personnel in the U.S. Office of Education is much encouraged over the progress being made at the grass root level in the evaluation of instructional programs. In previous years, for the most part, evaluation has not been playing a very important role in instructional programs, but under the Federal guidelines it has become necessary to evaluate programs to determine their effectiveness and furnish data on which to base decisions.

On the third day, the representative of industry, Edward C. Soistman, Manager, Program Planning for the Martin Marietta Corporation of Orlando, spoke on "Aero Space Management." This stimulating, illustrated lecture depicted very closely the parallel between what industry and what education should be doing in evaluating personnel and programs.

On the fourth day, Dr. Harry M. Sparks, President Murray State University, Murray, Kentucky, brought the thinking of the participants to an appropriate climax with his dynamic address, "Education For Imagination." He challenged the group to use evaluation as a means of striking out in an imaginative way.

The models developed by the working groups in Area A are presented to show what these several groups were able to accomplish. It is felt that they will serve as models for those who plan to develop instructional programs which can be adequately evaluated.

The participants of Area B spent their time working on the technical information which was to be used in evaluating the programs. Therefore, the report from this area constitutes an evaluation by the consultants and recorders of goals evaluated.

An overall evaluation was conducted by having the participants fill out a questionnaire which permitted them to rate how well each of them had attained the twelve principal objectives set forth by the State-Wide Planning Committee for the Seminar. Each participant at the first work session was asked to write down one or two personal behavioral objectives which he expected to attain during the Seminar. At the last session he was asked to rate himself on how well he succeeded. Opportunity was given for general comments and for making recommendations for future planning.

CHAPTER II

Addresses of Speakers

Two of the speakers, the first one and the last one, left with us texts of their addresses. They are reproduced as given. Mr. Jaeger did not have a prepared speech and his remarks centered around current opinions and events which were particularly of interest at the moment. Mr. Soistman's talk was an illustrated lecture which did not lend itself to inclusion in this report.

I. EVALUATION AND THE TWO-PARTY MONOLOGUE

Jarvis Barnes

Why the association of evaluation with the two-party monologue; with a long speech uttered by one person as if it were uttered by two parties? For the next several minutes, think of the evaluative program in your school system in terms of the extent that it is a long speech, figuratively speaking, carried on with a facade of two parties. Or maybe this is not the situation in your school system; but let's examine the evaluative process and, while doing so, suppose you mentally compare the process in your school system with some of the points which will be discussed. In other words, the multi-dimensional aspects of the process will be explored in the context of the consideration that evaluation is a system, consisting of components which in practice many times do not receive sufficient attention.

Your attention is now directed to four components of evaluation. These are: (1) the status evaluation has attained, (2) a few principles of evaluation, (3) evaluation as a subsystem within a school system, and (4) cost effectiveness. It is hoped that during this week, as attention is given to your specific assignments, that the points that are discussed will receive your attention.

How many of you would have included an evaluation section in your Title III proposals if you had not been requested to do so by the manual of instructions? I dare say very few would have; in fact, an examination of the evaluation sections of the proposals submitted before the manual of instructions specifically requested the inclusion of such a section shows that only casual, sketchy attention was given to evaluation. Now, even after rather specific instructions concerning the inclusion of such a section, a vast number of the proposals, I am informed, do not have exemplary procedures; for many are still sketchy, while many others state in more words than necessary that the program will be evaluated by some outside agency. This latter arrangement seems to indicate that the local school system feels it has clearly and sufficiently taken care of the matter, when, in fact, there appears no provision for some of the more basic contributions which should be made during the evaluative process.

Do the decision-makers of your school system view evaluation as a luxury? If you were required to reduce the appropriations requested in a proposal, would you tend to reduce the allocations for evaluation; that is, if the amount included in the budget were sufficient to allow for a reduction and if you thought the negotiators would be prone to accept the reduction? Would you not wish to reduce the appropriation because of the value that the findings generally have in terms of the decision-making process, in terms of continuance or modification of the program? Do the decision-makers of your school system really believe that evaluation procedures are worth the time and effort required? What do the teachers think about evaluation? Are both teachers and decision-makers involved in the process? If not, why spend money for it? To what extent is the evaluative process becoming an integral part of the operational and instructional processes? Do you pursue the evaluative process slavishly, keeping in mind the stated objectives and not varying a little from the predetermined procedures? Do the people in your system view the process as one that is mandated? Is there latitude in the procedure to design concentrated, small studies, which in the current vernacular are called mini-evaluation projects, conducted either by individual teachers or by small groups of

teachers? Don't you think that one objective should be to encourage teachers to pursue creatively their tasks, to think of new strategies, and to design as objectively as possible (perhaps with technical assistance) procedures to measure systematically the effectiveness of their new ideas? Let's not forget that the forces of the natural, physical sciences are more easily dealt with than those of the social sciences, that the human organism is the most complex organism, and that human power is more potent than nuclear power. Consequently, instructional strategies should be organized according to what is known about learning, and educators should strive to subordinate personal experiences to scientific methods as a basis for decision-making. Education will mature greatly as this procedure is developed. The status of evaluation in your system will increase as each decision-maker of the local school system increases his use of scientific methods for reaching decisions.

Now let's turn to the second consideration, some principles of evaluation. If we can agree that evaluation might best be viewed as a forerunner to the adoption and dissemination of change, then the first basic principle becomes, in the form of a question: how does the school system view the results; what is the mechanism for change that exists within the system; and how prone to change is the system? Thus one benefit of evaluation lies in developing theories and in advancing knowledge. Actualizing this benefit is most difficult to do, particularly since education is slow in developing those theoretical foundations which are needed for building a science. Furthermore, developing theories and advancing knowledge are difficult because quick answers to practical problems are rarely obtained. Nevertheless, progress has been made by such organizations as the National Institute of Mental Health (NIMH) and the Agricultural Department, because they have viewed research and evaluation as a means of advancing knowledge. However, only within the last three years has education seriously (and relatively extensively) used evaluation and research as a basis of operation. One significant hampering element, of course, has been the social nature of the institution.

Consideration of the first principle, using the results of evaluation to develop theories and to advance knowledge, leads to the second principle, which may be stated as a question: does evaluation foster the unification of curriculum development and instruction? Are the classrooms used too exclusively as miniature societies, which they certainly are not? Do teachers follow curricular guides as if they are the ultimate word? Do teachers feel or believe that they are the ones who determine the curriculum and not some indefinite "they," usually referring to central office personnel? Do teachers develop instructional strategies after planning rather carefully in accordance with the characteristics of the learner, the objectives to be accomplished, and the available media in the form of human and material resources? Do teachers use evaluation as a means of promoting the earning-learning concept, to take the pupils into the environment of reality? To what extent have the evaluative plans been designed with the purpose of serving as guides for such decisions, as compared to the plans serving ultimately only as a means to compare something with something else, which will end with stating whether or not there is a statistically significant difference? If evaluation is to promote the unification of curriculum development and instruction, teachers will need training in viewing the role of evaluation in this relationship. Furthermore, this training can be accomplished quite well while on the job during either formal or informal seminars, workshops, and courses. How many of the systems represented here have conducted well-thought-out seminars concerning evaluation and its contributions to curriculum development, instruction, and, above all, to decision-making? Aren't behavioral changes promoted as one perceives the relationships between his needs and the knowledge or information which is available to him? Wouldn't seminars concerning the contributions of evaluation promote changes in the behavior of teachers, which would, in turn, produce more effective teaching as the teachers discard impressionistic information and rely more on scientific information about education?

A third and last principle which will be stressed is that evaluation in an educational setting should be both objective and humanistic. For example, stating objectives in behavioral

terms is important. In fact, the act of stating objectives in any terms is important, for it helps one to recognize some of the elements which are involved in the accomplishment of the stated objectives. However, objectives should be stated, whenever possible, in measurable or observable terms. Furthermore, including objectives stated in humanistic terms is also important. For example, I recently visited a project that is destined to be successful, from both "hard" data and humanistic data points of view. But let's examine why it is destined to be successful humanistically. The espirit de corps, the wholesome attitudes, the interest, the professional zeal and competence, and the desire to operate an exemplary project permeate the actions of the staff. Compare this state of affairs to another situation in which two teachers who were participating in the same program were rather equally qualified; that is, professionally and in terms of service. However, one teacher assumed the position with the pupils that the undertaking was going to be interesting, challenging, and intriguing to them — that they would have fun. On the other hand, the second teacher assumed the position that "well class, since 'they' say that we've got to do it, let's go ahead and get it over; all of you will find things that you won't enjoy and won't like doing, but I guess that's part of life." Which of the two groups of pupils do you think will exhibit the more promising results? Obtaining "hard" data is important in the evaluative process, but do not forget to obtain the humanistic data also, and do not forget to take a look at the results from both of these approaches. You will want to remember this principle as you pursue your assignments this week, and you will want to attempt to supplement the "hard" data with the humanistic data. Such a process will more nearly reflect the realities of your operational situation.

Now before going into the third component of this discussion, let's pause and restate the three principles which already have been discussed:

1. Use the results of evaluation to develop theories and to advance knowledge;
2. Use the evaluative process to promote the unification of curriculum development and instruction; and

3. Include in the evaluative process both objective and humanistic data.

Now as we pause to reflect on these principles, and as we try to apply them to what is being done within each school system represented here, do you believe that you have been carrying on a two-party monologue within your school system? Have you had feedback from other significant persons and elements, or have you (as the one responsible for evaluation) been plodding your lonely way toward determining whether or not some comparisons are statistically significant? Of course, this determination is fundamentally important, but it is not sufficient. Possibly as the next component of this discussion is explored, there will be a fuller realization that accomplishing the technical aspects of evaluation is just the beginning of developing more than a two-party monologue within a school system.

Accordingly, let's begin the discussion of this next component of evaluation by asking, "How is education perceived in your school system?" Is it perceived as a reservoir of knowledge or as a means for developing the cognitive, affective, and psychomotor domains and their related skills? Is it perceived as a process for accumulating facts, or is it perceived as a process for developing conceptual and logical thinking related rather directly to real situations? How does one develop evaluative plans for a project without considering the objectives of the educative process as set forth by the hierarchy of the school system? Are the objectives of the proposed project in accord with the objectives of the school system? If they are not, then one must question the longevity even of a successful program. In other words, the climate in which a project operates must be in accord with the climate of the school system, or else serious consideration must be given to the means by which a new and promising idea might survive in an alien climate. Hence, the objectives of the "special projects" (in quotes for reference purposes only and not for operational purposes) must be in accord with the objectives of the local educational agency. One criterion which might be applied here to each and every one of the projects now in operation is: "Do the evaluative procedures stem from the objectives and philosophy of the school system?" Are the

objectives and philosophy of the special projects in accord with those of the school system? Are the evaluative procedures live and vibrant within the system, and are they well known by significant individuals? Remember that the administrators, teachers, and parents make it possible for evaluation to occur. Furthermore, they are the ones who produce the results of the program or produce the elements for change. Hence, the evaluative process interprets the finidings in relation to the theories and practices which are set forth by the superintendent, the principals, the supervisors, the teachers, the general public, and the board members. Evaluation, therefore, must be viewed as a subsystem within a system; for it is not a process that occurs in isolation. While the technological aspects of evaluation are some of the beginning and basic aspects, the realistic aspects come to light as evaluation is considered to be a subsystem within a larger system.

When recognition is given to evaluation as a subsystem, an additional dimension must be described in the process: that is, the effects of the institution on the obtained results, the effects of the bureaucracy, the effects of the setting, and the pluralistic nature of the climate. In fact, consideration of these effects should be incorporated in the plans for accomplishing the objectives of the project. Which is more important in your school system: administration or instruction? In a given school setting, which is more influential? For example, does the curriculum influence the type of schedule, or does the scheduling process influence the organization for instruction? If the former, why aren't there more instructional organizations based on flexible scheduling, team work among faculties, and utilization of differentiated staffs? Research from many sources today certainly indicates that these are desirable and will provide for more effective programs. Furthermore, how many of you have included in your evaluative procedures an examination of the administrative activities by which the project is operated, including their influences and their relationships? Have you recognized that the operation of a particular project which is somewhat different from the "regular" activities is really a subsystem within the school system? Have you attempted to look at the schools in which the program occurs as a group, or subsystem of schools, or at the class-

rooms in which the program occurs as a group, or subsystem of classrooms?

One approach to recognizing that evaluation is really a subsystem is employed whenever teachers are encouraged and assisted in designing and in operating their own mini-research activities. Wouldn't a school system profit if it were to adopt the slogan, "each teacher an action researcher"? In essence the slogan implies that each teacher doubts he is conducting the instructional process in accordance with the best thinking, that he wants to be more creative, that he wants to be more innovative, and that he wants to take off the old and put on the new, that is, if the new offers a promise of improvement. Has the bureaucracy provided ways and means for teachers to become research oriented? Does the current evaluative program of your school system promote this goal? Have the teachers been given time to learn and time to try out new ideas? When each individual teacher independently of the evaluative process prescribed by the institution carries out his own evaluative project, the procedure does provide a fundamental approach to get him to change his strategies in accordance with documented data. Designers of proposals must remember that many teachers of a school system do not readily accept the cleverly voiced objectives in a proposal, particularly when these objectives have been designed primarily to get the money and not to bring about basic changes in the instructional process. Furthermore, teachers frequently prefer to try their own brands of innovations, and not just to accept other ready-made brands. Consequently, classrooms ideally should be scenes of mini-evaluations, of mini-trials (m-i-n-i), and of new ideas. Incorporate the part into the whole; in other words, a project evaluation will be stronger and more meaningful when it incorporates the effects of the bureaucracy on the program and, vice versa, when it incorporates ideas for improving instruction which have been initiated by the classroom teachers. Possibly as this dimension becomes a part of the evaluative process, one might be more assured that evaluation will remain as a necessary process; for when the Federal government no longer requires programs to be evaluated, educators in general will still have such a requirement.

The fourth and final topic of discussion is that of cost effectiveness: the process of determining the most economical procedure per unit of specified change as, for example, per unit of change in achievement. The determination of appropriate cost effective procedures is in its infancy; however, educational scientists should vigorously pursue the development of these procedures in order to cope with the rapidly approaching question of the extent to which private enterprise should appropriately become involved in the educational process. Undoubtedly, much discussion will ensue during the next few years as the role of private enterprise is delineated in education. Possibly, a basic rule might be: a particular task should be done by either a private or a public organization in accordance with which one can accomplish the task most effectively at the most economical cost. Accordingly, we might pause here to ask ourselves, would private enterprise have accepted and incorporated within its practices some of the practices which we in education now regard with respect, even without a sound basis? Would private enterprise have endorsed unilateral contracts, as are now used in education? For example, when contracts are made with teachers, what specifically is offered in return for a year's salary? Doesn't the teacher promise to do the best that he can? How frequently do boards of education attempt to determine whether or not the teacher's portion of the contract has been fulfilled? Why should there be compulsory pupil attendance laws? How many pupils would teachers have if each teacher were paid on the basis of the voluntary attendance of the pupils and on the interest-holding power of the instructional strategies? Would many teachers have to change their procedures in order to have pupils in their classes? Furthermore, how many of you have successful projects which can be shown to be successful when cost effectiveness is applied? Shouldn't we wish to present for approval by boards of education projects which would give the most profitable yield for the least cost? But how many of us do this? And yet, there is not a single one of us who does not desire that our projects be the most effective in production and in cost. Therefore, why not incorporate in the evaluation program not only the statistical procedures for determining whether or not the program has been statistically different,

but also for determining whether or not it has been statistically different in terms of yield and in terms of cost? Quite possibly, a program might be statistically significant in yield but not in cost. For example, if what we do were based on what we know and if we only have a certain amount of money to spend in the school system, why not delete the twelfth grade and add either a kindergarten or a more extensive preschool program, extending even to the prenatal stage? Do you have data to show that greater good per dollar might come from such a move? Certainly, such a move would be contrary to tradition, which is part of the institution and the bureaucracy to which reference was made earlier.

And finally, school systems are having their tasks further delineated as far as evaluation is concerned. You are to be commended for the collective efforts exhibited by your presence and by your programs. All of you are pushing ahead in a well-thought-out way to meet the challenges presented by the evaluative process. Accordingly, during this week, as you go about your specific assignments, will you give some thought to the following:

1. Using the results of the evaluation to develop theories and to advance knowledge;
2. Using the evaluative process to promote the unification of curriculum development and instruction;
3. Including in the evaluative process both objective and humanistic data;
4. Analyzing the extent to which the evaluative process is a subsystem within the school system; and
5. Incorporating cost effectiveness as an integral part of the evaluative process.

Reflection on these concepts will possibly help to clarify the extent to which evaluation in your local situation is a two-party monologue. Why not make it a multi-party, a multi-directional communication network?

II. EDUCATION FOR IMAGINATION

Harry M. Sparks

Anyone who has studied the habits of man or the mores of a society recognizes the fact that we react like a pendulum. No matter what the process, in our anxiety to get it in operation our zeal carries us too far, and then, we try to correct our error, as reactionaries, we go too far to the opposite extreme. This was evident in our adoption of the concepts of of progressive education and some of our recent reactions toward it.

It might be better to liken this entire situation to a clock, since the pendulum, which represents only a small part of our people swings, and the clock, representing the greater mass, just sits there going around in circles and getting nowhere.

In recognition of this tendency of thoughtful man to react against extreme movements of the pendulum, and since we are now over-emphasizing science and technology in our society, I am advocating Education For Imagination.

Please do not connect these remarks with the reactionaries who believe we should return to the arts of liberation of the nineteenth century, and that we need Charles Elliott's cure for a sick high school curriculum. It's like using asafetida as a means of keeping away a disease in an era when innoculation has been proven effective.

We are living in the childhood of the age of science. More than one-half of the people that have ever been called scientists are still alive. Today we generally accept only those things we can measure specifically as being science. These we call physical science. We are generally discrediting the scientific findings in human relationships because of the many variables which we cannot eliminate, and we are evolving a concept that says, "if it isn't measurable, it doesn't exist and if it varies, it is not worthy of consideration."

Thus we have created a concept of "technology for technology's sake" and we have carried this development to the extent that it threatens to destroy its creators (maybe I should say its discoverers since these relationships of natural phenomena have always been here, and we have just recently discovered them).

In the nineteenth century the arts of liberation boys were riding high and science had the status which a trades course holds in education today. They had to give the courses special names to make them pallitable: natural history for biology and natural philosophy for physics. Today it is an honor to have a B.S. degree even though B.S.'s are made to march behind A. B.'s at graduation time.

Today some men of intellectual courage are challenging the status of science. Jacques Barzum in his *House of Intellect* challenges the use of the scientific method and deplores the fact that social science has provided means whereby man can make mistakes without risk of suffering for them because of the establishment of organized philanthropy and the newly developed forms of social insurance.

In the series of articles in the *Saturday Evening Post*, "Adventures of the Mind," Sir Herbert Read, Professor of Fine Arts at the University of Edinburgh, Scotland, in "Art And Life" implied that art and crime are closely related; the aptitudes, impulses, and tendencies toward destruction and creativity are the same. He quoted Tolstoi whose great novel of *War and Peace* was based on this concept. He states that the illness of the world has increased with the progress of technology. "We pull levers and push buttons and have ceased to make things with our hands." This is happening not only to the muscles and nerves but also to the creative processes of the mind. Mr. Read deplores the fact that pragmatic learning has made skill the understanding and control of machines rather than doing something.

When man is deprived of participating in the creative process, he directs his abilities to destruction. Today technical idealism is dominating the world and overwhelming all forms of moral idealism. It is deplorable that higher education

aspires to a mathematical perfection in which sensuous experience is no longer admitted as evidence.

We need to bring into balance the two forms of intelligence: thought and feeling. This cannot be attained by surrendering to conformity or withdrawing into ivory towers.

Today more imagination needs to be brought into education. Too frequently pupils are guided by experiences in science about which the outcome has been known for centuries. In our English classes we emphasize the analysis of stories and poems but are given little opportunity to write. In art we are given a recipe for the carrying through or culmination of an art process rather than providing a creative experience. Even in our government, precedent is considered to be the primary basis of thought rather than social need. In education we often try to work out formulas for curricular improvement and receipes for methods. We frequently teach the cyclical theory of history which emphasizes that man and nations go through the same process of rise and fall that the nations of the past experienced thus making a man a victim rather than the creator of historical experience. Such approaches to science, literature, art, and government leads us in paths through the same errors we have experienced in the past.

Man needs stability in something in order to maintain his sanity. He can find it only in expanded ideals, in expanded concepts of God and social service, in methods (science and reflection), and in principles (do unto others, freedom and responsibility). They are not to be found in rules, facts, and prejudice. (All Indians march single file; at least the one I saw did).

Benjamin Franklin, who started the Academy in Philadelphia, was educated through an apprentice system. His Academy made provisions for the educational needs of Colonial America, but was later distorted and formalized by the Academicians and had to be replaced with the high school.

Thomas Edison, the inventor of the incandescent light, was dismissed from the eighth grade as a dullard.

Henry Ford, who made the great contribution of mass production of automobiles, was not a product of formal education. Two young bicycle merchants invented the airplane while physicists were appearing before the Langley Committee of the United States Congress and declaring that an appropriation should not be granted because it was scientifically impossible to fly a heavier-than-air craft.

If Franklin Delano Roosevelt had been an educated scientist, he would not have sponsored the appropriation of funds which led to the creation of the atomic bomb and the development of modern atomic science.

As a college student, I was taught that it was impossible to destroy an atom. We were told that if the hydrogen atom were to be broken down that the explosion would detonate the streams, rivers, and oceans and the world would be destroyed. I made an "A" in that course and I would not have advised Congress to have made such an appropriation.

Don Garlit, a mechanic, developed a dragster that achieved a speed of over 186 m.p.h. within a quarter of a mile in 8.4 (eight and four tenths) seconds and disproved a law of acceleration which physicists have defended for years. If he had been taught physics by some teachers, he never would have tried it.

Is the only hope with the rebel? Maybe this explains the beatnik and hippie.

What can we do?

Self direction is going to have to be restored to learning; it is the only learning which is real anyway.

After fundamentals are started, and even during the process of their instruction, we must teach people:

How to study,

How to do research,

How to write poetry,

How to read intelligently,

How to participate in government,

How to live and work in a changing society, and

How to be imaginative.

You say, you can't measure it. You or I may not be able to measure these elements objectively, but both students and teachers can appraise the progress made in these areas.

Success in our society depends upon one's will and effort as much as it does upon one's capacity and information. However since we are able to evaluate capacity and information more objectively, these two factors have become the primary bases for measuring success in school. Imaginative teachers are discovering new means of evaluating and reporting to parents and employers that reveal a student's developmental possibilities.

Education is a process of becoming, not arriving, but we often treat it as a terminal process circumscribed by academic degrees. This continuing process of learning is enhanced or limited by integrity, industry, determination, purpose,judgement, and personality; and effective evaluative procedures should consider these characteristics.

The teacher and the student must learn to work together creatively in strengthening these characteristics that are so vital to the success and happiness of every individual. If we can consciously involve our students in this process of improvement they will react responsibly and creatively and they will be able to participate constructively in social change instead of trying to escape from it.

If we do not meet this challenge, academic degrees will become like rancid vats of oil; worthless because of age, and not refined sufficiently for use in the machines of modern thought.

Let's wake up to reality. Let's make education the vital living thing that it must be to effectively serve the needs of every person.

Let's use some imagination.

I wrote a poem:

IMAGINATION

I love to walk with thee, imagination,

To take some broke stringed harp and let my fingers play
lazily between its empty bars, and bring forth tunes
more sweet, more clear than man e'er dreamed of,

To take an untipped brush and splash on old torn canvas
scenes that Raphel would joy to gaze on,

To speak with silver-tongued voice and sway men's minds
to deeds they ne'r thought they could attain,

To lead them up the rays of setting sun to light the stars
in such a way to cancel out the night,

Although I know I cannot do these things,

You lead me higher than I was before.

Reality is grand, but forever is enhanced,

Because I walked with thee, imagination.

CHAPTER III

Area A—Evaluative Models In Selected Areas

J. B. White, General Consultant

I. INTRODUCTION

In Area A, six work groups were organized to develop a design for evaluation in five subject areas: reading, self-concept, mathematics, communication skills and guidance. The groups were asked to select a specific program in their respective subject areas and do the following things:

1. Design a program so that it can be evaluated. This implies a clear statement of objectives in realistic and behavioral terms.
2. Describe the processes that might be followed and techniques for evaluating these processes.
3. Explore the types of data that might be assembled, how such data can be secured, instruments that are available or may be developed.
4. Design the project. Set up simple designs that can be carried out by staff members who are relatively unskilled in statistical procedures.
5. Explore ways of interpreting the results and writing a final report.
6. The emphasis throughout will be on evaluation not the subject matter per se.

These groups worked towards developing a model which might be of value to themselves and to others in planning an evaluation program for these subject areas. While the groups were given the same general instructions, each chose to approach its problem from a slightly different point of view, emphasizing the fact that evaluation may follow many procedures and designs. With some minor editorial changes,

the reports from these groups are reproduced as they were developed. It is obvious that the greatest value stemming from the work of these groups came to the individuals who shared in the thinking and development of these models; however, the models themselves are significant and interesting. They should be suggestive to teachers and administrators who face problems of developing evaluation programs in these or other subject areas.

II. MODEL FOR READING PROGRAM I

Esther Miles, Group Consultant

Staff Development: Visual Perception

I. Identification of need

A. Examination of achievement test results at the end of the first grade reveals that 60 per cent of first grade children are not achieving in reading commensurate with their abilities.

B. Since it has been demonstrated by the twenty-seven first grade studies conducted nationwide by the USOE that the most significant factor in pupil achievement in reading is the teacher, it follows that a staff development plan for improving the teaching skills of first grade teachers will, in turn, affect pupil achievement in a positive manner.

II. Educational goals of staff development proposal

A. Each of the twenty selected first grade teachers will show improvement in teaching skills through their participation in in-service training workshops once a month during the FY 1969. These teachers will be selected because of their *inability* to teach perceptual skills to first graders.

B. Given consultant help, each of the twenty first grade teachers will be able to teach these skills to perceptually disabled pupils.

III. Behavioral objectives of the proposal

A. Given a class of twenty five, six and seven year-old children, the teacher will be able to demonstrate the techniques of the Frostig Perceptual Development Program within the following limits and conditions:

1. *Diagnosis*

The teacher will administer and score the Frostig Test of Visual Perception according to instructions accompanying the tests.

2. *Instruction*

The teacher will:

a. Classify the children according to criteria developed for the following categories:
   (1) No perceptual instruction indicated
   (2) Visual-motor training indicated
   (3) Figure-ground training indicated
   (4) Perceptual constancy training indicated
   (5) Position in space training indicated
   (6) Spatial relations training indicated

b. Conduct an instructional program for the children in each category according to the instructions in the Frostig Manual.

B. At least sixteen out of twenty teachers (or 80%) will be judged proficient in order that the program may be ascertained successful.

IV. Situations in which presence or absence of desired behavior can be observed or recorded.

A. Each teacher shall present to the group of other teachers a tape recording of an actual test administration with a child, together with transparencies of the test pages showing the child's responses, the scoring of the responses, and the interpretation of test results.

B. Mid-term testing will be done to measure program effectiveness as a facet of *process,* as opposed to *product,* evaluation.

C. At the end of the instructional program, each teacher will administer the Frostig Test of Visual Perception and reclassify students according to III-2a.

D. Each teacher will present post-test results and reclassification of students to the teacher group.

V. Norm or standard to be used in measuring teacher proficiency in the use of the Frostig program

A. Each teacher shall be able to administer and score the Frostig Test accurately. No errors will be accepted. The judgment of the consultant will be the criterion for assessment of the teachers' proficiency in test administration, scoring, and interpretation.

B. The t-test of significance will be applied to children's pre-and post-test scores to determine teacher effectiveness in implementing the instructional program in perceptual development.

VI. Application of evaluation methods

A. Demonstrations of testing proficiency as outlined in IV-A shall be presented to the group after the consultant has had sufficient time to instruct the teachers in the techniques of test administration, scoring, and interpretation.

B. The pre-test (Frostig Test of Visual Perception) will be administered to the first grade children in the twenty selected classes during the third week of school.

The post-tests will be administered during the last week in January and again during the second week in May.

VII. Analysis of evaluative evidence

A. An analysis of the testing proficiency of the teachers will be made by the consultant according to the criterion in V-A.

B. An analysis of the significance of mean gain will be made. The .05 level of confidence will be accepted.

VIII. Statement of conclusions

A. A statement of conclusions will be made concerning the ability of teachers to administer, score, and interpret the Frostig Test.

B. A statement of conclusions will be made concerning the ability of teachers to implement the Frostig Program.

C. The results of the program will be further examined to determine its continuation in terms of these questions:

1 Is the program a successful model for other teacher training programs?

2. Is it economically feasible?

3. Is perceptual development an effective method of improving the teaching of reading?

III. MODEL OR READING PROGRAM II

John L. Spagnoli Group Consultant

EDITOR'S NOTE: The group working on this model developed an elaborate and complicated model which involved the process of identifying the reading needs, developing and evaluating an in-service education program for teachers to meet these needs, and identifying the success of the program in meeting the child's reading needs. Since this program was complicated and involved really three programs in one, the editor has attempted to abstract from it the suggestions for identifying the reading needs and evaluating the success of a program attempting to meet these needs.

Identifying the Reading Needs

Before the general and behavioral objectives can be established for a program in reading, it is essential to know the specific needs of the children concerned. The following steps are suggested:

1. Construct or choose an evaluative instrument which will measure specific skills in reading relative to the instructional level.
2. Administer such a test to children individually and use the responses to determine the reading needs.
3. Analyze and evaluate individually and collectively the

reading needs for the purpose of determining the general and behavioral objectives desired.

Using the information secured from the initial testing program, the teacher will develop a statement of general objectives which the data have indicated are needed. From this list of general objectives, a number of behavioral objectives will be designed specifically to meet the needs of the children concerned. At this point, the curriculum content and activities judged by the teacher to be necessary to attain these objectives will be planned.

Process Evaluation

During the instructional period, the teacher will administer at intervals certain tests to determine the progress of the children towards the behavioral skills established in the general and the specific objectives. This information will help to evaluate the instructional process and to make changes as the test results might indicate progress or lack of progress towards the specific objectives.

Product Evaluation

At the end of the instructional period, the teacher will administer such tests as can be constructed locally or secured commercially to measure the extent of progress that has been made by each child towards the objectives outlined. Locally constructed school tests, because of their close proximity to the individual needs of the child, should be given high priority; however, standardized tests should be used so that comparisons with national and regional norms can be made.

Test results can be used for the following purposes:

1. To determine needs that should be met by subsequent instructional programs.
2. To provide a basis for planning and developing the curriculum.
3. To determine the extent of enrichment programs that should be provided.
4. To indicate weaknesses in the instructional program and to serve as a basis for in-service training designed to improve teacher efficiency.
5. To provide a basis for revising future programs of reading instruction.

IV. MODEL FOR ARITHMETIC PROGRAM

Donald Altieri, Group Consultant

A Comparison Of Two Methods Of Teaching Regrouping In Addition At The Primary Level

I. *Introduction*

Computational skills are important as an integral part of the total education of youth. There is a need to evaluate various existing and new strategies for teaching computational skills. In the new mathematics programs, there is a definite need to reevaluate the methods of teaching basic skills.

The purpose of this study is to compare two methods of introducing regrouping in addition to primary students. It is understood that results in different situations may vary because of variables other than the methods used. However, if one method seems to result in the students' better understanding of the given process, it would be worthwhile to continue classroom research in in this area.

II. *Problem*

When primary students have had no formal instruction in regrouping, does the supplementary use of Cuisenaire Rods significantly affect their ability to regroup in addition?

III. *General Objective*

To teach regrouping in addition.

IV. Behavioral Objectives

The student will be able to:

A. Rewrite the numbers correctly using place value notation when given expanded numerals in unnatural groupings.

B. Add numbers whose sums require the process of regrouping without the use of manipulative materials.

V. *Definition of Terms*

A. Unnatural Grouping — any grouping where smaller units can be expressed as larger units.

Example: 1 ten plus 14 ones equal
2 tens plus ? ones

B. Cuisenaire Rods—colored rods of graduated length which may be used to represent our number system.

C. Formal Instruction — a planned lesson which introduces a given concept.

D. Regrouping — number groupings which may be expressed as smaller or larger units.

E. Manipulation — handling objects.

F. Objects — Any materials which may be handled by the learner.

VI. *Hypotheses*

A. There will be no significant increase in the ability of students to regroup in addition when Cuisenaire Rods are used as supplementary aids.

B. There will be no significant difference between the control and experimental groups in the performance of:

1. rewriting the numbers correctly using place value notation when given expanded numerals in unnatural groupings.
2. adding numbers whose sums require the process of regrouping without the use of manipulative materials.

VII. *Design*

The Posttest Only Control Group Design was adopted because both groups are assumed to be at the same level of development and are being formally introduced to regrouping for the first time. Furthermore, the need for the administration of a pretest to determine bias is compensated for by a random selection of the groups.

VIII. *Procedures*

A. Two primary groups of like abilities are taught by the same instructor. One group is designated as a control group, the other as the experimental group.

B. In addition to the instructional materials available to both groups, the experimental group will utilize Cuisenaire Rods.

C. The study will involve instructional periods of the same duration for three weeks.

D. Process evaluation is an on-going evaluation of activities for the purpose of determining the degree to which the behavioral objectives have been satisfied. This evaluative process will be designed to test measurable performance.

The instructor will use an informal inventory checklist of observable behaviors which could be observed by having the student or students manipulate objects, work exercises on the board or paper, and by responding verbally to teacher questions. This inventory checklist will consist of two categories: satisfactory or unsatisfactory.

E. A test specifically evaluating the process of regrouping in addition will be obtained from the test data bank of the Computer center. (If such a data bank is not available, a teacher-made test may be used.)

IX. *Analysis of Data*

A. Comparison will be made in achievement gain by equating the experimental group against the control group using the mean of each sample group as the base and applying the t-statistic as a measure of significance of differences between groups.

Specifically, the performance of the two classes will be compared using the t-statistic. The .05 level of confidence will be criterion of the hypotheses in question.

B. Graphic comparison of the performance will be obtained by plotting frequency of responses against raw scores obtained from posttest data. For comparison purposes both control and experimental data will be included on the same graph.

C. The informal inventory checklist will be examined to note trends in relation to behavior attainment with regard to stated behavioral objectives.

X. *Signifiance*

The significance of the project would be determined by examining the findings and relating them in general terms to the problem and stated objectives. Specifically, it is hoped that this project will add to the knowledge of how to teach regrouping in arithmetic and generally provide a model for further investigations.

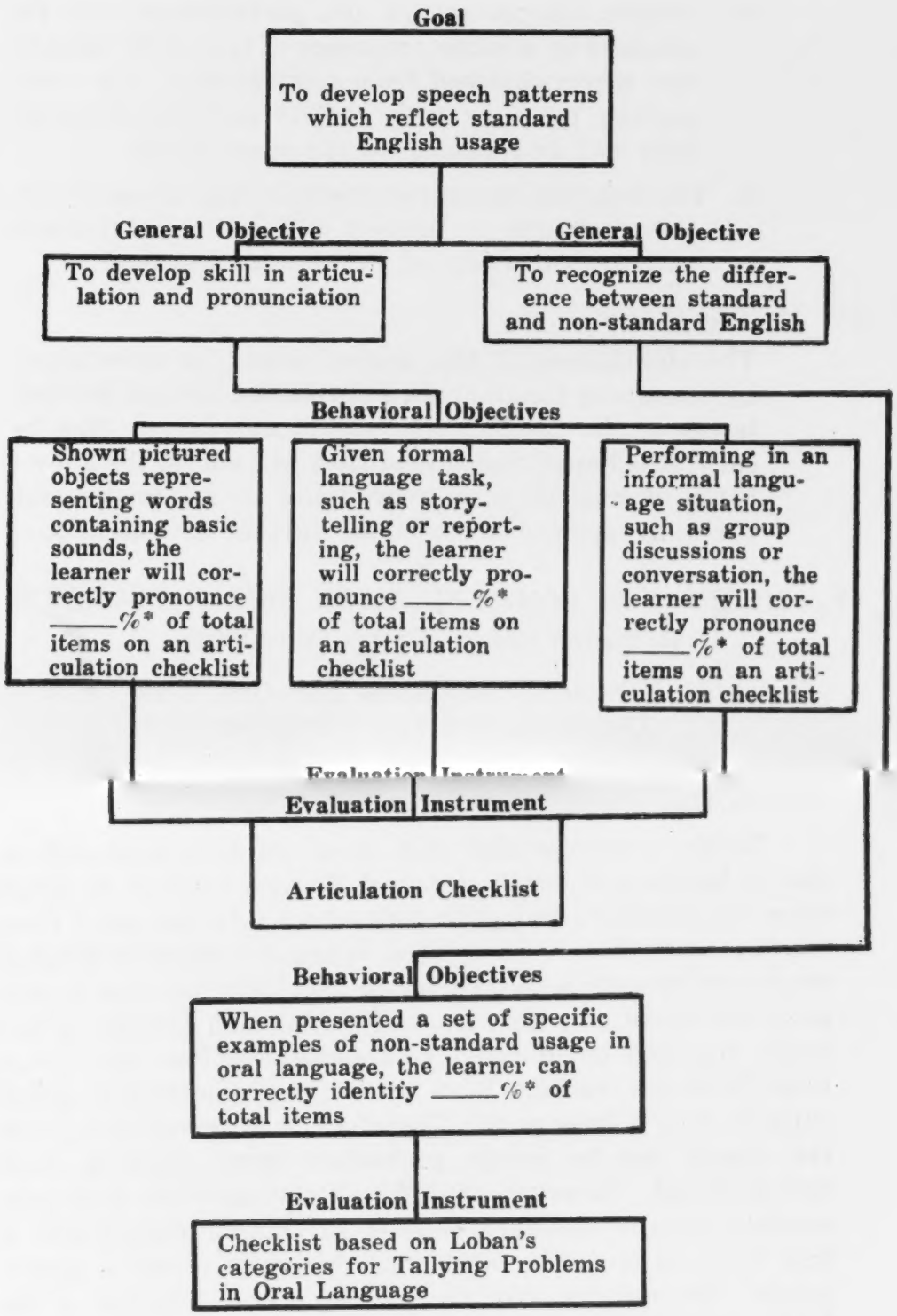
V. MODEL FOR COMMUNICATION SKILLS PROGRAM

Ruthellen Crews, Group Consultant

Suggested Behavioral Objectives For One Area In A Communication Skills Program

Introduction

Teachers acknowledge that some children have difficulties in learning to speak standard English because it differs from the language they bring to school. At the same time, teachers know that children need to acquire standard English usage-not because it is superior in itself, but because society penalizes those who do not use it. Standard English is not easily acquired by children who sense rejection and disapproval from the teacher. Such teacher-response simply makes children stop talking at all. Therefore, it is agreed that children should not be totally prohibited from speaking their native dialect. However, the school's responsibility is to help children acquire standard English as another dialect and to help them to recognize when such language usage is appropriate. Recognizing this, the following are examples of behavioral objectives which might be included in a program whose overall goal is to develop speech patterns which reflect standard English usage.


Goal
To develop speech patterns which reflect standard English usage
General Objective
To develop skill in articulation and pronunciation
General Objective
To recognize the difference between standard and non-standard English
Behavioral Objectives
Shown pictured objects representing words containing basic sounds, the learner will correctly pronounce _____%* of total items on an articulation checklist
Given formal language task, such as story-telling or reporting, the learner will correctly pronounce _____%* of total items on an articulation checklist
Performing in an informal language situation, such as group discussions or conversation, the learner will correctly pronounce _____%* of total items on an articulation checklist
Evaluation Instrument
Articulation Checklist
Behavioral Objectives
When presented a set of specific examples of non-standard usage in oral language, the learner can correctly identify _____%* of total items
Evaluation Instrument
Checklist based on Loban's categories for Tallying Problems in Oral Language

*Acceptable % and time required to achieve this skill to be determined by teacher who writes objective.

To elaborate on the first behavioral objective as stated "shown pictured objects representing words containing basic sounds, the learner will correctly pronounce ____%* of total items on an articulation checklist" — steps for evaluating progress could follow this procedure:

1. Identify specific problems (*i.e.*, non-standard) in articulation and pronunciation which are common to children in a specific classroom or school. For example, problems in use of consonant sounds of non-reading children may be determined by showing each child a set of individual picture cards which would elicit orally a word containing the consonant sounds in various positions in the word. For this purpose, pictures representing these words could be used:

Consonant Sound	*Initial*	*Middle*	*Final*
b	boat	cabbage	tub
d	dog	puddle	hand
f	father	muffin	knife
g	girl	wagon	frog
h	house	behind	
k	key	turkey	book
l	lamb	collar	ball
m	mouse	hammer	farm
n	nose	pencil	barn
ng		singer	ring
p	pig	apple	cap
r	rabbit	shirt	car
s	sun	postman	horse
sh	ship	machine	dish
ch	chicken	teacher	match
t	tie	mitten	gate
th	thumb	nothing	tooth
v	vine	river	stove
w	wood	twins	
wh	white		
y	yellow	barnyard	
z	zoo	magazine	rose
zh		treasure	garage
j	jacket	engine	page

2. Once the problems have been identified, a series of instructional activities should be planned which would provide for sequentially presenting the articulation and pronounciation of standard English.
3. A series of short evaluative devices consisting perhaps of words drawn from the above word list could be presented at planned intervals to measure progress made by each child.

A similar checklist of words containing vowel sounds could be made.

Progress in the second and third behavioral objectives — "given formal language task, such as storytelling or reporting, the learner will correctly pronounce ____%* of total items on an articulation checklist" and "performing in an informal language situation, such as group discussions or conversation, the learner will correctly pronounce ____%* of total items on an articulation checklist" — could be measured by taping each child at planned intervals as he is participating in a formal language activity (storytelling or reporting) and an informal language activity (conversation or dramatic play) and evaluating on the basis of changes in production of consonant sounds.

As children begin to make marked progress, tapes should be made of as many activities as possible for peer and self-evaluation as well as teacher evaluation.

To determined areas in oral language which children are unable to discriminate as standard and nonstandard English as required in evaluating the behavioral objective as stated — "when presented a set of specific examples of nonstandard usage in oral language, the learner can correctly identify ____%* of total items" — the teacher presents orally (perhaps by tape) two phrases or sentences illustrating the standard and nonstandard forms and asks the children, "Which is standard?" Examples of sentences which deal with verb problems and which might be included in the evaluative device could be as follows:

He say he is going home.
He says he is going home.

The boy don't look happy.
The boy doesn't look happy.

We have to see it because he want to see it.
We have to see it because he wants to see it.

My mother look at television a lot.
My mother looks at television a lot.

One girl have a basket on her bike.
One girl has a basket on her bike.

They runs down the street.
They run down the street.

The two little girls looks at the little boy.
The two little girls look at the little boy.

She ask him while they walks home from the movie.
She asks him while they walk home from the movie.

We likes to ride our bikes in the park.
We like to ride our bikes in the park.

You turns that machine on too fast.
You turn that machine on too fast.

I sees it.
I see it.

I is going outside.
I am going outside.

We is the best ones.
We are the best ones.

I thought you was going to ask me that.
I thought you were going to ask me that.

They was three girls.
There were three girls.

Here is two dogs.
Here are two dogs.

He happy.
He is happy.

That girl my friend.
That girl is my friend.

These examples are taken from *Loban's Problems In Oral English* (National Council of Teachers of English, 1966). He has identified nine categories for use in tallying problems in oral language. For further elaboration and specific examples for each category refer to his publication.

VI. MODEL FOR GUIDANCE PROGRAM

Robert D. Myrick, Group Consultant

Guidance

I. *General Summary*

The evaluation of guidance is a complex but essential undertaking. Many factors contribute to helping an individual, both inside and outside of school. To separate and assess the influence of the school guidance services from that of other factors constitutes one of the major aspects of guidance evaluation.

Simply defined, evaluation consists of making systematic judgments of the relative effectiveness with which goals are attained in relation to specified standards or procedures. Systematic program evaluation may be done in relation to professional standards. In this case, a formal checklist or survey may be utilized in order to compare program with recommended standards. Or, systematic evaluation may be done in relation to specific guidance procedures that are thought to lead to the attainment of individual and professional goals.

The evaluation of guidance has been approached many ways. Whatever approach is used, however, three component elements are inherent in any comprehensive evaluation. First, the objectives of the program must be stated in observable behavioral terms. That is, the anticipated outcomes of whatever is being done must be formulated in such a manner that they can be verified. Second, the procedures or methods for obtaining these objectives must be clearly established and operationally defined. Third, data collection activities must be

organized so that accurate evidence may be available for final analysis and appraisal. At least three different evaluation models are available: (a) The Survey Approach, (b) The Case Study Approach, (c) The Experimental Approach.

The Survey Approach is probably the most frequently used appraisal or evaluation model in guidance. Predetermined criteria should be selected for inventory. This method is limited as to whether student behavior has been significantly affected. The survey approach tends to emphasize the availability of activities, staff, facilities, and programs. The lack of experimental validation, difficulty in inferring casual relationships and sampling errors are among the major difficulties which bias survey data. This in turn has led guidance personnel to more rigorous and specific kinds of evaluation models.

The Case Study Approach is becoming more popular as a result of the new emphasis in studies of behavior modification (Krumboltz, Bijou, etc.). As opposed to case conferences, where the data is selected after the fact and then presented for study, the case study approach features the establishment of baseline data prior to guidance or counseling activities. Following the activities, or concurrent with them, observation continues and is plotted according to a frequency polygon. Although time consuming, it emphasizes the study of the individual and his growth. It avoids the massing effects of other models.

The Experimental Approach is fast becoming the most popular model for evaluating the effectiveness of guidance programs. It involves carefully planned steps to study one or more groups of individuals in terms of one or more variables. Scientific methods including a predetermined sequence of procedures are used. This model has many forms, but the most appropriate are those that require control and experimental groups. Experimental control is used in order to ascertain whether change or gain by the procedure to be investigated can be attributed to the procedure or method of treatment rather than to chance alone. Major designs to be considered, are: (a) The Before and After Design; (b) The After-Only Design. N. L. Gage's book, *Handbook of Research on Teach-*

ing, Chapter 5, is considered a major reference for experimental and quasi-experimental models.

II. *Seminar Activities*

In general, the group attending the Seminar at Orlando covered the following general outline:

A. Operational Procedures and Definition of Problems
   1. Statement of general guidance objectives, specific objectives, problem of study, definition of terms used
   2. Development of behavioral objectives

B. Three Evaluation Models: (1) Survey, (2) Case Study, and (3) Experimental
   1. Advantages and limitations of each approach
   2. Sampling procedures, collection of data methods

C. Criterion Measures
   1. Relationship to guidance procedures—revelant evaluation

D. The Analysis of Data — Distribution Free Statistical Tests
   1. Correlated Data
      a. Sign Test
      b. Wilcoxon's Matched-Pairs Signed-Ranks Test
   2. Uncorrelated Data
      a. Mann-Whitney U Test

E. How to write a report

F. Mock Department Meeting was part of final meetings. All members of the group acted as staff and one interest area was identified and considered for evaluation. Steps of deliniation and delimiting experimental design and related problems of appraising the final results were discussed.

III. *Evaluation Model*

The result of the final meeting was a sample model which was concerned with the evaluation of the effects of group counseling on underachieving students. This sample model follows:

The Effects of Group Counseling on Underachieving Ninth Grade Students

I. *Introduction and Statement of the Problem*

It is the purpose of this study to investigate the effects of group counseling on underachieving ninth grade students in "X" high school.

A. *Importance of the Study*

(A paragraph stating the need for this study is assumed.)

B. *Definitions of Terms Used*

1. *Underachieving.* This terms refers to students whose total battery scores on the Comprehensive Tests of Basic Skills and California Test of Mental Maturity show a significant difference between their obtained grade equivalent and anticipated achievement grade equivalent.
2. *Group Counseling.* This refers to a group of five students and a state certified counselor who will meet in fifty-minute sessions, once a week for nine weeks during a school year. The procedures in the sessions will be based upon the recommended group counseling process under the July, 1968 EDPA proposal as submitted by twelve Florida counties under the direction of the State Department of Education.

II. Review of Related Literature.

(This section would contain a pertinent review which would support the rationale for this study, including professional literature and previous attempts of evaluation and research at the local level.)

III. *Population and Sample*

Population. The population for this study consists of 100 underachieving students (70 boys and 30 girls) from the ninth grade class of "X" high school in "Y" county. The total ninth grade population consists of 600 students.

Sample. For purposes of investigation, forty students

will be randomly drawn from the 100 students in the population. These forty students will in turn be randomly assigned to either an experimental or control group.

IV. *Experimental Design and Procedures*

A. *Experimental Design.* The before and after design will be used as described in N. L. Gage's *Handbook of Research on Teaching,* Chapter 5. This design would require criterion measures administered prior to group counseling and again at the end of counseling. This allows not only for a comparison between the experimental group (20 students) and the control group (20 students), but also a comparison of individuals based on their own pre and post data.

B. *Experimental Procedures.* Experimental procedures as follow will be used:

1. Using a table of random numbers draw the sample (40 students) from the population of underachievers, based on the major criterion listed below.
2. Using a table of random numbers assign (20 students) of the sample to four groups of five each, all of whom will receive group counseling. The remaining twenty students will serve as a control.
3. Randomly assign two of the experimental groups to each of the two counselors in "X" high school.
4. Obtain the pre-group counseling data from the criteria listed below:
   a. *Major Criterion.* CTBS Comprehensive Test of Basic Skills and California Test of Mental Maturity-Short Form. (Administered once each spring of the school year.)
   b. *Supplementary Criteria.* Grade Point Average (collected once every nine weeks) Attendance (Number of days attended each nine weeks) Teacher Rating Scale-five-point scale would rate each student on: (a) parti-

cipation in class, and (b) completed assignments.

5. Group counseling will begin the first week of the second nine weeks period. Groups will meet once a week for nine weeks.
6. With the exception of the major criterion measure, all post data will be collected following the end of group counseling and again at the end of the fourth nine weeks.
7. Pre and post data will be analyzed according the Mann-Whitney U test and other appropriate statistical procedures.

V. *Report of the Findings*

A report of the findings will be written, which will include:

1. the problem
2. the design and procedures
3. the criterion measures used
4. the findings and
5. the summary, conclusions and implications

VII. MODEL FOR SELF-CONCEPT PROGRAM

Richard C. Rank, Group Consultant

A Project To Improve Self-Concept of Disadvantaged Second Grade Children.

I. *Statement of Need*

Children tend to behave in ways which fit what they believe about themselves. Because of social conditions of family, peer culture, and educational practices, children receive much negative comment about their behavior and their personal worth. Conversely, our cultural habits more often prevent us from freely and naturally complimenting others for job well done or for just being a good person.

This emphasis on "negatives" and the relative absence of "positives" causes children to have unrealistically low estimates of their basic worth as people, their social competence,

their power, their intellect, and sometimes their physical appearance. Such low self-estimates frequently result in defensive, destructive behavior, low motivation, and withdrawal.

The schools can help correct this low self-estimate of many children without "additional cost to the taxpayer" through more intensive teacher use of sincere praise, verbal reward of success, and statements building feelings of personal worth.

These teacher techniques should be systemically tried out, and resulting changes in student self-concept and behavior should be carefully assessed. If results are significant in value, such practices should be encouraged on a wide scale.

II. *General Objective*

The broad goal of this project is to help students feel and think better of themselves, their capabilities and personal worth, resulting in more effective and growth-producing behaviors.

The following outline is intended to provide one possible example of how to evaluate a student self-concept improvement project carried out in an elementary school. The evaluation procedures for the project encompass both *process* evaluation and measurement of project *outcomes*. Since the conference focus is on evaluation procedures, the skelatal outline of the project treatment is not a well developed model. It only provides a procedure, or idea to be evaluated and should not be used in other ways.

Emphasis is placed, however, upon carefully selected *behavioral outcomes*, as well as a direct measure of self-concept (the California Personality Inventory.) Such behavioral outcomes were selected as additional valid indicators of positive self-concept.

III. *Behavioral Objectives*

Behavioral objectives will be stated in two ways:

A. A change toward higher estimate of self as indicated by higher scores on the California Personality Inventory, and

B. An increase in rate of selected specific behaviors:

1. *classroom participation,* as evidenced by asking questions, answering questions, volunteering information, and participation in group activities;
2. *effective problem-solving* behaviors such as, purposeful appraisal of problem before beginning the task, choosing a strategy, trying other strategies, and perseverance; and
3. *effective social interaction,* as demonstrated by willingness to share, to take turns, to help others, to make suggestions, and to accept directions or suggestions from others.

IV. *Project Outline and Procedures*

In lieu of a complete resume of project design and procedures, a brief outline will set the stage for sample evaluation procedures.

Theory and research in "self-concept" suggests that more expression of positive feelings toward students by others, more recognition and reward for successful student efforts, programming of student tasks to insure success, and increased frequency of positive physical touch or contact by others, all serve to increase students' positive feelings about themselves.

In order to carry out the treatment process for the "Positive Feedback Project," teachers will be given additional training in:

A. making a variety of positive statements to students about their performance, social behavior and personal worth;

B. use of more physical contact demonstrating positive regard and affection;

C. development of special tasks and situations in which student success is insured; and,

D. recognition methods, reward systems and incentives for successful performance.

As part of the treatment, teachers will be asked to tabulate, (See Table I) day by day, the frequency with which they employ such methods in the school setting. The treatment phase will last for one report period. Evaluation of outcomes will be done through pre- and post-testing, as outlined.

This project is presented in a demonstration design which will only show amount of change in self-concept and selected behaviors. An experimental design comparing the effectiveness of such a program with existing procedures, or establishing its effectiveness beyond chance, will require a revised design employing additional control or alternate treatment groups.

TABLE I

Teacher *Jones* Observer *Webb*

Time Span 9:00-10:00 A.M.

TEACHER BEHAVIOR CHECKLIST
(for use during treatment phase)

Date	Method			
	Positive Statments About Student Performance	Physical Contact	Use of Success Tasks	Use of Reward Systems and Incentives
Oct. 15	~~1111~~ 1	~~1111~~ 1	1	11
Oct. 16	11	~~1111~~ 111	11	111
Oct. 17	~~1111~~ ~~1111~~	11	111	1

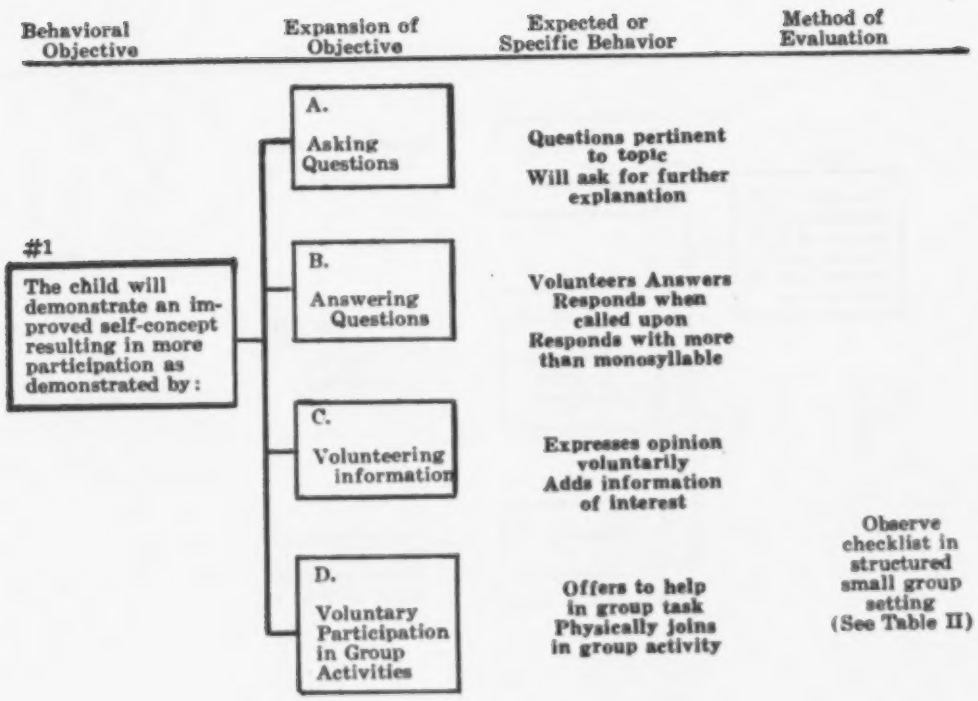

Behavioral Objective
Expansion of Objective
Expected or Specific Behavior
Method of Evaluation
#1
The child will demonstrate an improved self-concept resulting in more participation as demonstrated by:
A. Asking Questions
Questions pertinent to topic
Will ask for further explanation
B. Answering Questions
Volunteers Answers
Responds when called upon
Responds with more than monosyllable
C. Volunteering information
Expresses opinion voluntarily
Adds information of interest
D. Voluntary Participation in Group Activities
Offers to help in group task
Physically joins in group activity
Observe checklist in structured small group setting (See Table II)

TABLE II

CHECKLIST FOR BEHAVIORAL OBJECTIVE #1

Participation

Date_____

Group_____

Teacher_____

Pre-Post Measure of Student Participation Patterns

BEHAVIORS

Name	Asks Questions	Volunteers to Answer Questions	Volunteers Information	Voluntary Group Participation	Total
1. Mary Smith	111		11		5
2. John Jones	1				5
3. Bill Johnson	~~1111~~			~~1111~~ 1	12
4. Etc.					

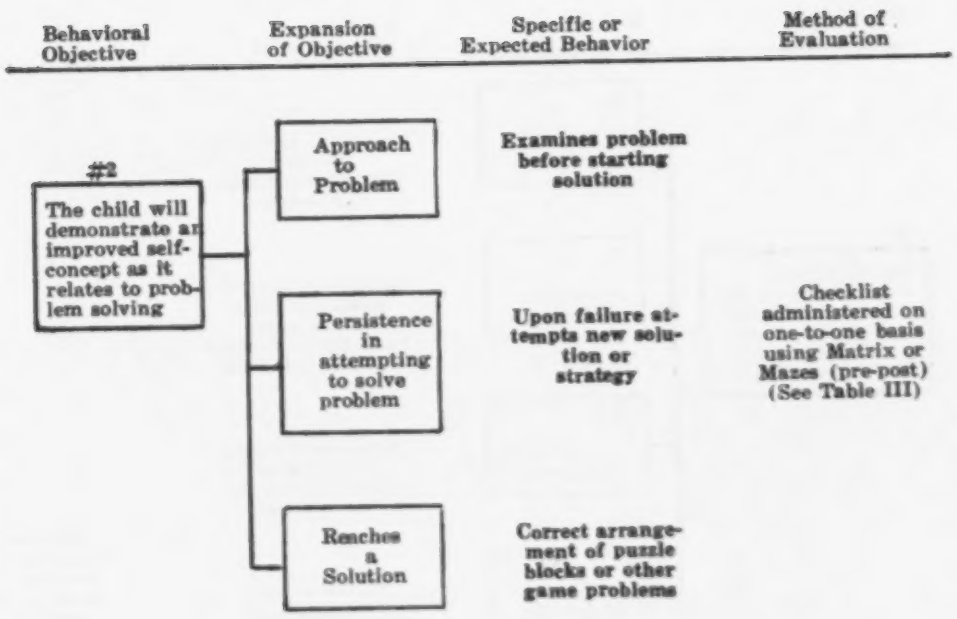

Behavioral Objective
Expansion of Objective
Specific or Expected Behavior
Method of Evaluation
#2
The child will demonstrate an improved self-concept as it relates to problem solving
Approach to Problem
Examines problem before starting solution
Persistence in attempting to solve problem
Upon failure attempts new solution or strategy
Checklist administered on one-to-one basis using Matrix or Mazes (pre-post) (See Table III)
Reaches a Solution
Correct arrangement of puzzle blocks or other game problems

TABLE III

CHECKLIST FOR BEHAVIORAL OBJECTIVE #2

Student Problem-Solving Behavior

Date_____

Group_____

Teacher_____

Pre-, Post Measure of Student Problem Solving Behavior

BEHAVIORS

Name	Time Spent Examining Problem Before Starting	Frequency of Trials of New Approaches	Reaches Solution
1. Mary Smith	30 sec.	111	Yes
2. John Jones	10 sec.	1	No
3. Bill Johnson	20 sec.	11	Yes
4. Etc.,			

Behavioral Objective	Expression of Objective	Specific or Expected Behavior	Method of Evaluation
#3 The child will demonstrate an improved self-concept by more peer group interaction	Willingness to Participate	Volunteers suggestions Volunteers to play part of family Physically remains with group	Observer checklist used in small group setting in which students are instructed to role play a large family preparing to go on a picnic, or a school class going to the zoo. (See Table IV)
	Sharing, or in other ways putting another's need before one's own	Statements such as "let me help you!" Lending paper or other props "You be the Daddy. I can be the little boy"	
	Accepting judgment or direction from others	Carrying out a task assigned by a peer Becoming quiet when asked	
	Attempts to direct or in other ways promote goals or the group	Accepting a role assigned by a peer Statements like "let's do it this way or "We need to have a mother" or "We can't go on a picnic without a lunch"	
	Showing positive regard for others	Hugging or other expression of caring or liking Statements to others using words like "love" or "like" Smiling at others	

TABLE IV

CHECKLIST FOR BEHAVIORAL OBJECTIVE #3
Social Interaction

Date_____

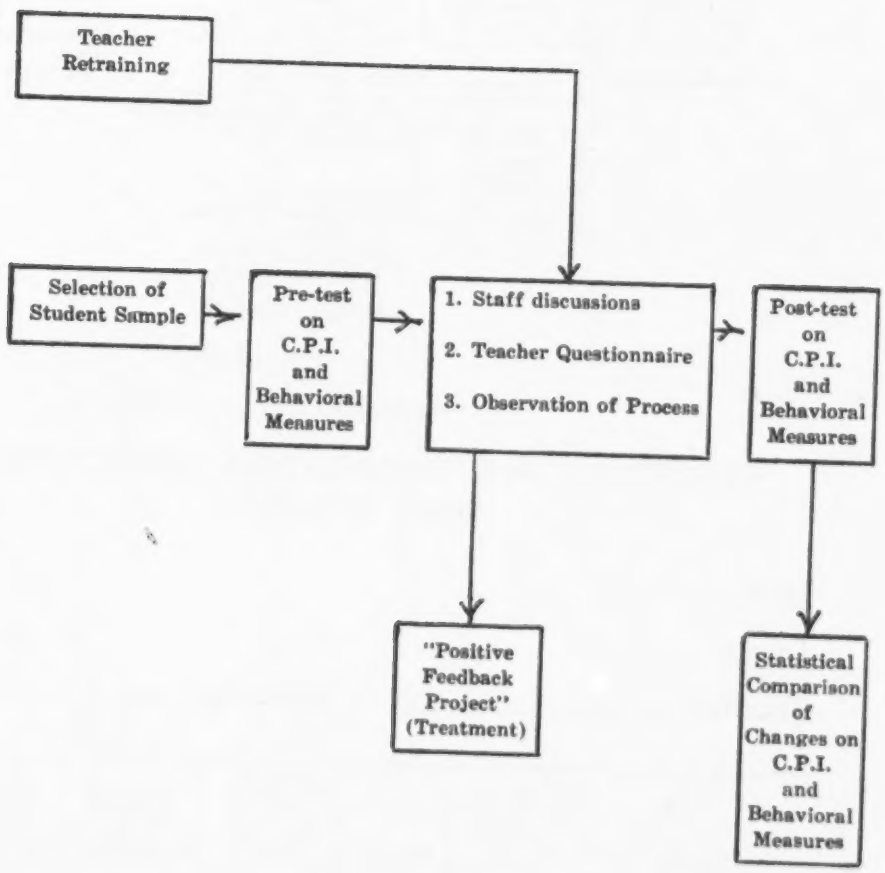
Group_____

Teacher_____ Observer_____

NAME	Voluntarily Participates	Sharing	Accepts Judgment or Direction from others	Attempts to Direct or Promote Group Goals	Shows Positive Regard for others	Total
1. Mary Smith	11	111	~~1111~~ 11	111	1	16
2. Bob Jones	1	1	11	11	1	7
3. Bill Johnson	~~1111~~	111	11		1	11
4. Etc.,						

PROJECT FLOW CHART

PROCESS


Teacher Retraining
Selection of Student Sample
Pre-test on C.P.I. and Behavioral Measures
1. Staff discussions
2. Teacher Questionnaire
3. Observation of Process
Post-test on C.P.I. and Behavioral Measures
"Positive Feedback Project" (Treatment)
Statistical Comparison of Changes on C.P.I. and Behavioral Measures

CHAPTER IV

Area B — Technical Evaluation Information

H. A. Curtis, General Consultant

I. INTRODUCTION

Several days prior to the Seminar, the Consultants in Area B were furnished a topical outline. The memorandum accompanying it stated that it should be considered a resource document intended to suggest the content suitable for this Seminar. The memorandum also pointed out that each Consultant should expect to consider only those topics which seemed to be of value to his particular group and omit those which were already understood, too abstract, or irrelevant.

Immediately prior to the opening of the Seminar, the Consultants met to go over the outline, clarify the meaning of some of the terms, and to exchange suggestions for the presentation of technical material to the Seminar participants. The topical outline follows.

II. TENTATIVE TOPICAL OUTLINE — AREA B

I. OBJECTIVES

- A. Translating global objectives into observable objectives.
- B. The role of the evaluator in pinning down the realistic objectives of projects.
- C. Relationships with teachers during and following the process of defining objectives, etc.

II. INSTRUMENTATION

- A. Using tests included in the school's testing program.
  1. Re-scheduling to obtain pre-post measures.
  2. Utilizing part-scores where applicable.

3. Re-scoring of selected items which are valid measures of specific objectives.

B. Building short tests for specific objectives:
   1. Sources of items (adaptations of text materials, teacher's quizzes, "adapting" items from other tests, etc.)
   2. Characteristics of good test items
   3. Test item format.
   4. Directions to pupils (and teachers.)
   5. Importance of try-out and revision.
   6. Scoring and scoring systems (raw, percentile, standard, etc.)

C. Subjective data gathering devices:
   1. Reporting behaviors vs. opinions.
   2. Illustrations of kinds of data that may be gathered.
   3. Principles of constructing such devices.
   4. Quality of data-frequency counts, generally.

D. Teacher observations and judgments.
   1. Role in evaluative studies
   2. Anecdotal data
   3. Treatments of such data

E. School records.
   1. Attendance, absences, tardiness.
   2. Disciplinary reports.
   3. Health records, etc.

III. DATA GATHERING DESIGNS

A. Coverage of all pupils; vs.

B. Sampling and spiraling;
   1. Difference between individual diagnosis and assessing group change.
   2. Sampling principles
   3. Spiraling as a means of measuring 3 or 4 objectives in minimum time. Also as a means of getting more frequent readings for time line analyses.

IV. SOURCES OF TEST INFORMATION

A. Use of standard references

V. SCORING AND PROCESSING

A Machine scorable answer sheets.

B. Over-Printing.

C. Simple scoring machines.

D. Layouts for quick hand-scoring.

VI. DATA TYPES AND ADMISSABLE PROCESSES

A. Nominal (classification data)

B. Ordinal (ranks, percentiles, scale positions)

C. Interval (Standard scores)

D. Ratio (fiscal and physical measurements)

VII. DESIGNING EVALUATIONS

A. Try to get across the purpose(s) of good design through chalk talk or other informal presentation. In the process develop basic concepts and vocabulary, including:

1. Treatment variables, independent variables.
2. Outcome variables, criterion variables, dependent variables.
3. Sources of invalidity, contamination, confounding which cause the evaluator to claim too much, too little, attribute results to wrong causes, or generalize to wrong situations or populations. (See: Gage. *Handbook of Research on Teaching,* Chapter 5; (Campbell & Stanley.)

VIII. DESIGNS

Eight designs will be provided as handout materials. These have been prepared by Dr. Westbrook. Note the reference to "Guidelines." This is a publication of USOE intended to aid in the evaluation of Title I Projects. It presumably was sent to all counties and has been used by the participants. Note also "Sources of Invalidity" section in most presentations. The explanation of these terms should be covered in Topic VII, immediately above.

IX. STATISTICS

A. Descriptive function of statistics

1. Frequency counts, percentages, percentiles.

2. Graphical presentations (Histograms, Frequency polygons, Cum f, Scattergrams)
3. Central tendencies (Mode, Median, Mean)
4. Measures of dispersion (Range, Q, Standard deviation)
5. Correlation (Product moment, Rank order)
6. Exercises — Data sheets will be prepared.

B. Inferential function of statistics.
1. A plain explanation of "inference," and illustrations of when we need to use it.
2. A plain explanation of "error," its sources, and why it prevents the acceptance of obtained relationships at face value.
3. A plain graphical explanation of "standard error," and the rationale for its use in the t-test. Normal distributions.
4. A plain explanation of "statistical significance" and of "confidence levels."
5. Calculations
   a. $\bar{\sigma}$ m
   b. $\bar{\sigma}$ r
   c. $\bar{\sigma}$ p
   d. $\bar{\sigma}$ diff $(M_1 - M_2)$ (uncorrelated)
6. Using t-table, determine significance of differences between uncorrelated means, o's and p's. Explain effects of correlated measures.
7. Interpret results — let participants explain the difference between such statments as
   a. The mean score of the E group was 80; the C group 50; and the difference between means 30.
   b. The difference between the means was significant at the .05 confidence level.
8. Other tests of significance.
   a. F-test — Explain when it would be used instead of t.
   b. Specific tests for special distributions such as several tables for non-parametric tests.

X. NON-PARAMETRIC STATISTICS

A. Utility of distribution-free, small sample statistics.

B. Simplicity of many such tests — for instance, the sign test.

C. Limitations — (loss of power; often limited to small numbers; usually can handle but a single variable.)

D. Useful non-parametric techniques.

1. Chi-Square
2. Sign test
3. Run test
4. Others, if appropriate.

XI. TECHNIQUES FOR HANDLING SEVERAL VARIABLES IN ONE EXPERIMENT

Possibly explain the basic idea of ANOV; regression analysis; and factor analysis to let participants know that complex designs can be used.

XII. ELEMENTS OF GOOD WRITTEN REPORTS

Importance of the audience. Difference between formal report and report to teachers and to lay groups.

III. EVALUATION — AREA B

The consultants and recorders were supplied a form which provided for their respective evaluations of the work of the Seminar and for their suggestions for future seminars or similar training programs. The form was subdivided into four principal sections. The responses for each section will be summarized, serially.

I. *Ratings of Principal Topics of the Syllabus*

This section listed the twelve principal topics in the syllabus and requested each respondent to:

(a) cross out all topics which were not taken up in his group, and

(b) to rate each topic on a 1-5 scale (unsuitable-very good)

The number of consultants reporting that they did not take up the topic is reported in Table V in the column headed "Number Omitting." The median of the consultants' ratings and of the recorders' ratings are reported in the columns headed "Consultants' Ratings" and "Recorders' Ratings." The number of consultants rating a topic "2" or "1" is reported in the last column. No recorder accorded any topic a rating below "3."

The data in Table V indicate that five topics were omitted by one or another of the consultants and that no topic was omitted by two consultants. The recorders' reports indicated that fewer topics were omitted. The differences between the consultants' and recorders' reports can probably be accounted for by the fact that the consultants reported as "omits" the topics they did not specifically present, while the recorders excluded from the "omits" topics which were discussed informally in the groups.

TABLE V

SUMMARY OF CONSULTANTS' AND RECORDERS' RATINGS OF TOPICS

Topic	Number Omitting	Median Consultants' Rating	Median Recorders' Rating	No. Rating "2" or "1"
1. Objectives	1	4.5	5.0	
2. Instrumentation		5.0	5.0	
3. Data Gathering Designs		4.0	5.0	
4. Sources of Test Information	1	4.0	4.5	
5. Scoring and Processing		3.0	4.0	2
6. Data Types and Admissable Processes		3.0	5.0	1
7. Designing Evaluations		4.0	4.0	
8. Designs		4.5	5.0	
9. Statistics (Parametric)		4.0	5.0	
10. Statistics (Non-Parametric)	1	3.0	4.0	
11. Techniques for Handling Several Variables	1	3.0	4.5	1
12. Elements of Good Written Report	1	4.0	4.0	

A comparison of the medians of the consultants' and recorders' ratings reported in columns two and three show that on nine of the twelve topics, the recorders tended to rate more favorably than did the consultants. They tended to agree in their ratings of three of the topics. In no instance did the recorders rate topics lower than did the consultants. It is not possible to determine whether this quite systematic difference should be attributed to "generosity error" on the part of the recorders, or to an unrealistic aspiration level on the part of the consultants.

The entries in the last column report the number of consultants who rated topics "2" and "1" which indicates questionable suitability as topics for this Seminar. These responses, when coupled with the median ratings of the consultants, suggest that "Scoring and Processing," "Data Types and Admissable Processes," and "Techniques for Handling Several Variables" were of questionable value to the participants in this Seminar. "Non-Parametric Statistics" was not rated very favorably by the consultants, but none indicated that it should be omitted.

II. *Specific Items*

This section applied to specific items, concepts, skills, etc., included under the larger topic headings of the syllabus. Its purpose was to identify specific items that were:

(a) inappropriate, because most of the participants already had the competence represented,

(b) particularly appropriate because participants saw its relevance to their programs and were ready to learn,

(c) or were inappropriate because they were too abstract and/or participants perceived them to be unrelated to the programs for which they were responsible. The responses of the consultants and participants are summarized below.

Inappropriate, Because Participants Already Possessed the Competence Represented

No single item was identified by the consultants or by the recorders as belonging in this category. One consultant noted that about half of his group could start basic statistics at a level above zero, and others noted that there were items with which some members of their groups were competent, but there were no items with which all members were competent.

Particularly Appropriate Because Participants Saw Its Relevance to Their Programs and Were Ready to Learn

The responses to this request are difficult to interpret because of a general tendency to identify whole topics. The spacing on the report sheet was scant and possibly discouraged

2

the giving of detailed responses. A plausible interpretation may be that the respondents tended to find most, if not all, of the items under a topic appropriate. The data show that the topic most generally considered appropriate by consultants and recorders was "Designs." "Instrumentations" and "Parametric Statistics" were tied for second place in terms of frequency of mention. In addition, "Objectives," "Sources of Information," "Designing Evaluations" and "Data Gathering Designs" were each mentioned by one consultant and one or more recorders. One recorder endorsed all topics except "Objectives" and "Elements of a Good Written Report."
Seven specific items were identified. These were:

1. Administrative level at which responsibility for evaluation should be placed
2. Relationships with teachers during and following the process of defining objectives
3. Building short tests for specific purposes
4. "Spiraling
5. Meaning of kinds of variables
6. Graphical presentation of data
7. Understanding of statistical "error"

Inappropriate, Because They Are Too Abstract and/or Participants Perceive Them to be Unrelated to the Evaluations for Which They are Responsible

The majority of the consultants identified the advanced statistical techniques for the handling of several variables as being inappropriate for these participants. Non-Parametric statistics and inferential statistics when presented in the short time available were also mentioned by one or more consultants as being inappropriate. The comment of one consultant was to the effect that much seemed inappropriate to the participants, but not to him and suggested that the perception of the participants should be changed.

The responses of the recorders were limited. Two items were specifically identified by one recorder. These were "statistical calculations" and the use of the T-table. The writer's interpretation of the comment of one recorder is that

he felt that no part of the program was inappropriate because the participants needed basic knowledge, and not an introduction to evaluation using a process approach that would be applicable to all types of educational programs.

III. *Other Observations About This Seminar*

The purpose of this section was to give the consultants and recorders an opportunity to present any criticisms or suggestions that might be helpful or informative to the Steering Committee members.

The suggestion most frequently offered by consultants was that a pretest or other screening device be used to assign participants to groups more homogeneously. This would serve the unsophisticated by enabling the consultant to deal at greater length with basic concepts and skills. It would serve the most sophisticated members by permitting the establishment of at least one group that could deal with advanced topics and the newer developments in the area.

Other suggestions offered by the consultants were that participants should have an outline of the material to be covered and an appropriate text for their use throughout the Seminar; that a certificate of proficiency be awarded on the basis of a score earned on an examination: that task forces visit participants in home counties to assess the carry-over into practice; and that a less attractive locale might encourage night work and discourage night life.

The recorders were generally agreed that some plan for homogeneous grouping was needed, but were not explicit about the method to be used. The recorders seemed to favor a system in which the participants would elect the group with which they would work. The majority of the recorders stated that the work was too much for the time allotted and that a more informal atmosphere and more breaks were needed. One suggested the addition of program development to the content to be covered. One suggested that the technical concepts and skills need to be more closely related to practical applications. Some recorders complimented their consultants and others commented favorably about those who planned the program.

IV. *Suggestions For Future Training Situations*

The members of the committee who had planned this Seminar were involved officially or because of their professional interest in planning future training programs for people involved in evaluation. For this reason, the responses of the consultants and recorders were solicited to several specific questions and to an open-ended request for any further comments. Their answers and suggestions are summarized below.

Consultants' Suggestions

The consultants were unanimous in the opinion that in future situations a syllabus should be furnished all participants. The writer is of the opinion that they would favor a detailed syllabus or even specially prepared text material. Their judgment as to the size of the groups in future situations ranged from six to fifteen. Ten seemed to be about the number most would favor. Instructors for future programs should be experienced in the practical aspects of evaluation, capable of supplying practical illustrations and applications of the techniques presented, and generally able to relate well with public school people. Not all consultants expressed an opinion about the length of future training conferences. Those who did, stated or implied that more over-all time was needed, possibly a full week, but breaks should be scheduled into the working day.

The possibility was posed that regional conferences might be held and that the same participants might attend more than once per year. The consultants' suggestions ranged from two sessions to five sessions per year. One apparently polled his participants and reported that most of them preferred one annual state-wide Seminar. If repeated sessions were to be held, the suggested intervals between sessions ranged from one to six months. Those who suggested more than two sessions recommended shorter intervals between sessions. One suggested that better grouping at an annual session might obviate the need for repeated sessions.

The consultants' suggestions for topics in addition to those which seemed to them quite appropriate at this session varied from "none," or no response to six. The extended list was:

1. Basic measurement theory and descriptive statistics
2. Objectives and criterion measures
3. Design and inferential statistics
4. Introduction to computers
5. Advanced computer applications

One suggested that time be provided for consultation on participant's individual projects; the availability of commercial test displays, provided that the publisher's representatives not be permitted to make presentations in the working sessions.

Other suggestions included making time available for consultants to work together in planning and preparing materials; providing reference materials, reprints and bibliographies; and suitable texts for use by the participants. One suggested a well constructed test of broad objectives and implied that such test would focus the participants' attention upon the specified objectives and also serve as a classification instrument for grouping. One consultant suggested a plan by which two groups, A and B, could work in parallel and put their work together as a complete proposal at the end. At step 1, Group A would work on objectives and Group B on specific behaviors to be observed; at step 2, Group A would work on procedures and Group B on instrumentation; at step 3, Group A would work on content and Group B on statistical analysis; at the fourth and final step, the two groups would fit their work together to make a complete project. The writer wishes to note that this plan was discussed informally and it was pointed out that "A" and "B" might be two individuals from the same system or small groups working on a single project. One consultant questioned the realism of trying to provide the training needed in any manner other than prolonged and systematic study of the type ordinarily furnished on campus or possibly in organized field courses. Another suggested that persons not responsibly involved in evaluting programs be excluded from training sessions. (Approximately 54% of the attendants at this Seminar were not so involved.)

Recorders' Suggestions

The recorders were unanimous in their recommendation that a syllabus be supplied in future training sessions, but

one noted that it should be used only to give the participants an overview of the area to be covered. Three of the recorders recommended that the groups range in size from ten to fifteen and two from eight to ten or eleven.

The recorders varied in their perceptions of the specific competencies, other than subject matter, which the consultants should possess. One stated there were no special competencies required and one did not respond. The others suggested a dynamic personality, a knowledge of group dynamics, and an ability to present technical material in the vernacular of the participants.

The future conferences should be three or four days in length in the opinion of the recorders. One noted that they might well be shorter if a specific objective could be attained in that time.

The recorders appeared to be quite widely separated on points pertaining to regional conferences that might involve several sessions per year. One recorder thought that the state-wide conference was preferable to regional conferences, another that there should be only one conference per year in each region, two suggested two sessions in each region and one five sessions. If regional conferences were held, the recorders thought that the topics considered should relate quite closely to the projects of the participating counties. They suggested that specific uses of the conference content in the participants' projects be stressed; that individual county programs be analyzed; that an evaluation of evaluation reports be conducted; and that examples of subjective rating instruments be developed and considered.

The recorders offered a number of suggestions for improving the evalution competencies of public school people. They do not fit into any general classification scheme, so will be reported as unrelated suggestions. The need to narrow the range of participants within groups was again mentioned. Two suggestions were offered:

(a) Section on the basis of their sophistication in the area; or

(b) on the basis of the type of responsibility they had

for evaluation on the job. In the event that a series of conferences could replace the single annual conference, then each of these conferences should have narrower objectives or be designed to serve persons working in specified roles. The writer interprets several of the comments to mean that a number of the recorders felt that this conference attempted to cover too much for too many kinds of people and in the future both the range of the objectives and content, on the one hand, and the range in the sophistication and interests of the participants, on the other, should be limited.

The suggestion was again emphasized that "real projects" be used in the conference. Each participant should bring one of his own and actually execute the work at the conference. The writer is not clear about how the recorder making this suggestion thought this could be done in a conference, but he is clear on the point that the recorder thought this conference was too abstract. Other recorders clearly shared this view.

There were suggestions that the State Department furnish leadership in the field of instrument preparation, particularly validated attitude scales; that either the Universities or the State Department furnish multi-variate services (meaning not clear); that the Universities require systematic study of evaluation in the pre-service program of teacher preparation; and that those who attended this conference should, in turn, set up conferences in their local counties.

V. Details of Evaluation

A. Process of evaluation

1. Meetings with the project coordinator and participating teachers should be held weekly to discuss the procedures being used. Revisions in procedures may be necessary and should be made accordingly.

2. An anonymous questionnaire should be completed by teachers during the second and fifth weeks of the project in which they express concerns,

feelings, new ideas, changes or other aspects of the project. Project changes may be made based upon these teacher reactions.

B. Outcome evaluation

1. Feelings of personal worth (items from California Personality Inventory) will be used to provide a "self-concept score." These items should be administered approximately one week prior to, and during the week following, the treatment phase of the project. Average difference of pre- and post-scores of students can be used to find significance of score changes due to treatment.

2. Behavioral objectives (all to be measured before and after treatment within the same group structure)

   a. Class participation—Ten small groups of ten students will be shown pictures from the Level 2 Peabody Language Development Kit during three fifteen-minute activity periods the three days prior to the beginning of the project. Students are then asked to "ask or tell anything about the pictures." An observer should tabulate behavior of each student as shown on the Participation Behavior Checklist. Such observer, possibly interested mother, should be trained to use rating sheets by applying them to some students in regular classroom interaction.

   b. Problem-solving behavior—Individual students will be given the color puzzle block set and instructed to arrange the four blocks so that all four colors (green, blue, red and white) show on all sides of the block series. Times and frequencies should be recorded as per the Problem-Solving Checklist.

   c. Effective social interaction-Ten small groups of ten students will be asked to organize and assume roles of a large family getting ready

to go on a picnic. Fifteen minutes of interaction or all relevant behavior up to the assumption of family roles should be tabulated as per the Peer Group Interaction Checklist. On the post-test, role playing should be about the teacher, the teacher aide, the bus driver, and the children getting ready to go to the zoo.

VI. Statistical Analysis

In this project, a straight forward "eyeball" comparision of change in the frequency of desired behaviors from pre- to post-test provides some basis for deciding if the project has been worthwhile, and if the methods used in the project should be used on a wider basis. Small changes in frequencies should be viewed with caution, however.

For a complete and statistically sound analysis of data, the analysis of variance technique should be applied for, "between teacher," "between variable," and "between pre- and post-testing" variance. A "between student type" analysis can be done if students are grouped before analysis is done, based upon pre-test results on the C.P.I. scores.

CHAPTER V

Seminar Evaluation

I. EVALUATION OF SEMINAR OBJECTIVES

In this project, a straight forward "eyeball" comparison objectives and purposes for the Seminar that were set forth by the State-Wide Steering Committee. Participants were asked to rate how well they felt the Seminar contributed to their attaining of these objectives on a scale from one (slightly) to five (extensively). One hundred nine participants completed this checklist.

Table VI contains a list of these objectives followed by the median rating given each objective by the participants. The Seminar median rating for all objectives was 3.70. Those objectives whose medians fell below this overall median are followed by a "—," those above it by a "+".

A median rating was computed for each participant. The Seminar median of these is 3.747. Of the 52 participants whose median rating fell below the interval containing the overall median, 19 (or 36.5 per cent) were members of Group A and 33 (or 63.5 per cent) were of Group B. Therefore, a considerably large proportion of the ratings of participants in Group B fell below the Seminar median. A Chi Square test of these numbers showed that such a difference had a probability greater than .025, but less than .05.

Another comparison (Mann-Whitney U-test) made of the ratings of the participants in Group A *vs.* Group B showed a probability of .017 of this difference being due to chance alone. Therefore, there is a significant difference in the ratings given by participants in Group A in comparison to those given in Group B, tending to indicate that the objectives of the Seminar, as stated, were more appropriate to Group A which dealt with evaluation models than to Group B which dealt with technical evaluation information.

TABLE VI

SEMINAR EVALUATION CHECKLIST

1. Gained an increased awareness that evaluation is central to a sound educational program.	4.09	+
2. Gained a greater knowledge of the steps involved in the evaluation of an educational project.	3.86	+
3. Gained a greater understanding of how objectives grow out of needs.	3.59	—
4. Gained increased skill in writing objectives clearly and in terms of behavior that can be observed and measured.	3.64	—
5. Developed skill and understanding in developing model projects that can be evaluated properly.	3.63	—
6. Gained an increased awareness that a comprehensive evaluation should include areas related to the affective domain, and should include social and physical development.	3.51	—
7. Became familiar with a variety of instruments and devices that may be used to collect data about instructional programs.	3.54	—
8. Gained a clearer understanding of the criteria and procedures for determining the reliablity and validity of evaluative instruments.	3.77	+
9. Developed increased skill in analyzing evaluative data so that meaningful conclusions can be reached.	3.33	—
10. Became more familiar with a variety of methods for reporting results to various groups.	3.26	—
11. Learned about the more common pitfalls in the evaluation process.	3.83	+
12. Gained increased knowledge in devloping practical evaluation designs for pro jects and programs.	3.80	+

II. EVALUATION OF INDIVIDUAL OBJECTIVES

The individual objectives, as stated by the participants, were classified into eight general objectives in order that they might be studied and arrive at some understanding of how well the participants felt they had attained their own objectives.

The one objective which was listed by the greatest number of the participants (56), was "to gain a broader knowledge regarding the procedures for evaluating instructional programs." Using the same scale as before, the median score on this objective was 4.2 which indicates that the participants mentioning this objective were reasonably successful in attaining it.

The objective which received the highest rating of success (median 4.4) in the minds of the participants was "to learn how to develop an adequate design for an instructional program so that it can be effectively evaluated."

The next two objectives which had the same median score (4.3) were: (1) "to learn the techniques for writing behavioral objectives that are observable and measureable," and (2) "to gain a greater knowledge regarding the understanding and the measuring of a child's self-concept and attitudes and how these may be changed." These two objectives were uppermost in the minds of the members of the State-Wide Planning Committee. The high degree of attainment here indicates that the Seminar was reasonably successful.

The objectives which the participants felt they had not attained as well as those mentioned above were in descending order: (1) "to gain a knowledge of ways to interpret data gathered with the instruments used," (2) "to gain a broader knowledge of the methods and procedures for evaluating programs," (3) "to gain a greater knowledge of data gathering instruments," and (4) "to gain an understanding of how to write programs under the requirements of the EIE plan with particular reference to staff development."

III. GENERAL COMMENTS AND SUGGESTIONS

Participants were given the opportunity to make additional comments and suggestions as they desired. As would be expected, these were many and varied. An attempt has been made to categorize them under large general headings and then into sub-headings.

First reference is made to the general remarks made by participants regarding the success of the Seminar. Seventeen commented that the Seminar was very good or excellent. Twelve commented that the group leaders were excellent and knowledgeable. As was expected, several mentioned that the Seminar offered an opportunity to share ideas with one another, while several more commented that many suggestions useful in the writing of Federal projects were received.

There were others who commented about the Seminar schedule or its component parts. Three felt that the time devoted to the Seminar was too long, while at least one felt that the schedule was "too tight." At least one felt that the Seminar could be improved by holding the general sessions at a time other than at the lunch hour. At least two thought the general sessions could be held in the morning before the group meetings which also coincided with the feeling expressed by one person that the group work sessions were too long. At least two mentioned that the Seminar dates conflicted with the pre-school dates in their counties. At least two felt that the number in the work groups was too large and should be held to approximately ten.

The one suggestion mentioned by the most people had to do with the large disparity in experience and training of those participating within the working groups. Fifteen participants suggested that some way should be found to make the groups more homogenous. Two suggested that the grouping be done prior to the opening of the Seminar. Three persons felt that some type of outline for the Seminar should be sent to the participants before the opening date, while one felt that a bibliography would be helpful. One of the participants felt that teachers should be invited to attend. One person felt like he needed more help in the designing and evaluating

programs provided under the EIE fund. Two persons felt like the motel rates were too high.

These general comments and suggestions seem to indicate that most of the participants were very well satisfied with the Seminar, its location and program. Several helpful suggestions were made for improvements. No one made any comments or suggestions which were primarily of a negative nature.

The participants were asked to indicate whether they preferred the next Seminar, if held, be a a state-wide one as this one was, or whether there should be several regional Seminars. Ninety-three of the participants responded to this question with 75 per cent of them feeling that it should remain a state-wide Seminar, while 22 per cent felt that it should be regional in nature and 3 per cent felt that both types should be held.

CHAPTER VI

Recommendations For Follow-Up

At its post-seminar meeting on October 4, 1968, the State-Wide Steering Committee reviewed the report of the Evaluation Committee and considered its findings and suggestions. It was agreed that the Seminar had been very successful in helping the participants to attain to a rather high degree the objectives that had been agreed upon in advance. The importance of evaluating instructional programs is increasing greatly due to the requirements in Federal programs and now in the programs financed under the state EIE fund. It is sincerely hoped that the results of this Seminar will be felt back in the school systems where the participants are functioning this year.

The Steering Committee recommended again that another Seminar be held next summer following these recommendations:

1. The Seminar should be state-wide rather than regional in character.
2. The participants be invited on the same basis as has been over the last two years, *i.e.*, two from each county, from each state supported university and from each Catholic diocese.
3. That the Seminar be held in Orlando with the same personnel in charge.
4. The same general format to be used again with the general sessions occurring at the lunch period and the groups working in the mornings and afternoons.
5. That the first choice of time would be the second week in July, and the second choice would be the third week in June.
6. That it be held from noon on Tuesday through Friday noon.

7. That every effort be made to devise ways and means for assigning participants to the work groups in such a way that they will be more homogeneous in background, training and experience.
8. That the State Department of Education take the initiative in planning for this Seminar since the funds available for financing it would come from ESEA Title III funds allocated to the State.

It was felt by the members of the Steering Committee that these Seminars were meeting an outstanding need among Florida educators. It was also realized that many of the participants had been appointed recently to the positions for which they were poorly prepared. By attending the Seminar and gaining information and skills they were able to return home to do a much better job with confidence.

The work presented or reported herein was performed pursuant to Grant No. 68-06745-7 from the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

A limited number of copies may be secured from the Orange County Board of Public Instruction, P. O. Box 271, Orlando, Florida 32802.

Appendix

Seminar Leadership Personnel

CO-SPONSORS

Floyd T. Christian
State Superintendent
State Department of Education

James M. Higginbotham
County Superintendent
Orange County Public Schools

SPONSORS' REPRESENTATIVES

SEMINAR DIRECTOR
Leo H. Howell, Jr.
Coordinator, ESEA Title III
State Department of Education

SEMINAR COORDINATOR
Mills M. Lord, Jr.
Director of Federal Projects
Orange County Schools

SEMINAR ADVISORS

Dr. C. J. Martin
Regional Assistant Commissioner
HEW Regional Office
Atlanta, Georgia

Joseph Pukach, Director
Elementary and Secondary Educ.
Atlanta Regional Office
U.S. Office of Education
Atlanta, Georgia

Jon L. Stapleton, Director
Federal-State Relations
State Department of Education
Tallahassee, Florida

SPEAKERS

Dr. Jarvis Barnes
Assistant Superintendent for Research and Development
Atlanta Public Schools
Atlanta, Georgia

Richard Y. Jaeger
Program Planning and Evaluation
Bureau of Elementary and Secondary Education
U.S. Office of Education
Washington, D. C.

Edward C. Soistman, Manager
Progam Planning
Martin Marietta Corporation
Orlando, Florida

Dr. Harry M. Sparks, President
Murray State University
Murray, Kentucky

STATE-WIDE PLANNING COMMITTEE

Mrs. Lola T. Brady, Coordinator
In-Service Education
Orange County Schools
Orlando, Florida

James E. Carroll, Supervisor
Federal Project Planning
Orange County Schools
Orlando, Florida

Dr. John L. Cropp
Administrative Assistant
to Superintendent
Federal Projects
Sarasota County Schools
Sarasota, Florida

Dr. Hazen A. Curtis, Professor
Educational Research and
Testing
Florida State University
Tallahassee, Florida

Dr. Paul W. Fitzgerald, Director
Pupil Personnel
State Department of Education
Tallahassee, Florida

James W. Hosch, III
Research Associate
State Department of Education
Tallahassee, Florida

Leo H. Howell, Jr., Coordinator
ESEA Title III
State Department of Education
Tallahassee, Florida

W. L. Kitching, Director
Panhandle Area Cooperative
Project
Marianna, Florida

Mills M. Lord, Jr., Director
of Federal Projects
Orange County Schools
Orlando, Florida

Dr. H. H. McAshan, Director
Project IDEALS
Lakeshore Towers
Gainesville, Florida

Dr. Kenneth Schultz, Ass't. Dir.
Instructional Research,
Development and Evaluation
Dade County Schools
Miami, Florida

Phillip Sciortino, Supervisor
Catholic Schools
Diocese of St. Augustine

Dr. Bernice S. Scott
Educational Specialist
Duval County Schools
Jacksonville, Florida

Mrs. Lilly May Shaw
Administrator Federal Programs
Alachua County Schools
Gainesville, Florida

James R. Swanson
Executive Director
Division of Research
State Department of Education
Tallahassee, Florida

Royce B. Walden, Director
Research and Special Projects
Orange County Schools
Orlando, Florida

Dr. Annie W. Ward, Coordinator
Guidance and Testing
Volusia County Schools
Deland, Florida

Dr. J. B. White, Professor
Education and Executive
Secretary FERDC
University of Florida
Gainesville, Florida

GROUP LEADERS
AREA A-EVALUATIVE MODELS IN SELECTED AREAS

GENERAL CONSULTANT

Dr. J. B. White, Professor
Education and Executive
Secretary FERDC
University of Florida
Gainesville, Florida

Area	Group Consultant	Recorder
Reading	Mrs. Esther Miles Duval Reading Clinic Duval County Schools Jacksonville, Florida	Mrs. Peggy Scharfenberg Supervisor Perceptual Training Learning Disability Center Orange County Schools Orlando, Florida
Reading	John L. Spagnoli Coordinator of Reading Palm Beach County Schools West Palm Beach Florida	Alfred B. Klein Supervisor Reading Sarasota County Schools Sarasota, Florida
Mathematics and Science	P.K. Laboratory School Donald Altieri College of Education University of Florida Gainesville, Florida	Mrs. Idella J. Moss Supervisor Elementary Mathematics Sarasota County Schools Sarasota, Florida Andrew Conda Elementary Supervisor Hillsborough County School Tampa, Florida
Communication Skills	Dr. Ruthellen Crews Dept. of Elementary Education College of Education University of Florida Gainesville, Florida	Vilma Vega Pilot School Learning Specialist Hillsborough County Schools Tampa, Florida
Guidance	Dr. Robert Myrick Dept. Personnel Services College of Education University of Florida Gainesville, Florida	Dr. S. Allen Smith Chief Psychologist Duval County Schools Jacksonville, Florida
Self-Concept	Dr. Richard C. Rank Associate Professor College of Education Georgia State College Atlanta, Georgia	Mrs. Eunice T. Carter Coordinator HEARD Program Duval County Schools Jacksonville, Florida

AREA B-TECHNICAL EVALUATION INFORMATION

GENERAL CONSULTANT

Dr. Hazen A. Curtis, Director
Educational Research
School of Education
Florida State University
Tallahassee, Florida

Group Consultant
Dr. Louis Bashaw
Associate Professor
College of Education
University of Georgia
Athens, Georgia

Dr. Jacob G. Beard
Assistant Professor
School of Education
Florida State University
Tallahassee, Florida

Dr. Charles W. Bridges
Associate Professor
College of Education
University of Florida
Gainesville, Florida

Dr. Howard W. Stoker
Division of Research
School of Education
Florida State University
Tallahassee, Florida

Dr. Annie W. Ward
Coordinator of Guidance
and Testing
Volusia County Schools
Deland, Florida

Dr. Bert W. Westbrook
Assistant Professor of
Psychology
North Carolina State University
Raleigh, N. C.

Recorder
Dr. Shirley Griggs
Assistant Principal
King High School
Grosse Pointe, Michigan

Ray Hess
Case Worker
Learning Disability Center
Orange County Schools
Orlando, Florida

David J. Krasnosky
Evaluation Specialist
Escambia County Schools
Pensacola, Florida

Dr. Margaret W. Wood
Evaluation Specialist
Orange County Schools
Orlando, Florida

Roe M. Martin, Coordinator
Testing and Research
Polk County Schools
Winter Haven, Florida

Mrs. Phyllis Presley
Elementary Supervisor
Orange County Schools
Orlando, Florida

"Education in an educational setting should be both objective and humanistic."

Dr. Jarvis Barnes

"Reality is grand, but forever enhanced,

Because I walked with thee, imagination."

Dr. Harry M. Sparks